Performance Indicators

for

Mental Health Services

Values, Accountability, Evaluation, and Decision Support

Final Report of the Task Force on the Design of Performance

Indicators Derived from the MHSIP Content

June 5, 1993

TASK FORCE ON THE DESIGN OF PERFORMANCE

INDICATORS DERIVED FROM THE MHSIP CONTENT

Rand L. Baker
Oklahoma Department of Mental Health and Substance Abuse Services
P.O. Box 53277
Oklahoma, Oklahoma 73152
(405)271-8666
(405)271-7413 FAX

Judith Cook, Ph.D.
Thresholds Research and Training Center
2001 North Clybourn, Suite 302
Chicago, IL 60614
(312)348-5522
(312)348-4416 FAX

Trevor R. Hadley, Ph.D.
Department of Psychiatry
University of Pennsylvania
3600 Market St., Room 717
Philadelphia, Pa. 19104
(215)662-2886
(215)349-8715 FAX

Edna Kamis-Gould, Ph.D.
632 Bryn Mawr Ave.
Penn Valley, PA 19072
(215)667-3569
(215)667-8415 FAX

Walter Leginski, Ph.D. Division of Demonstration Programs, CMHS
5600 Fishers Lane
Room 7C-08
Rockville, MD 20857
(301)443-3706
(301)443-6349 FAX

Ted Lutterman Director of Computer Operations & Research
NASMHPD
1101 King Street, Suite 160
Alexandria, VA 22314
(703)739-9333
(703)548-9517 FAX

Jack Morgenstern, M.D.
VP for Medical Affairs
Hallmark Health Care
P.O. Box 723049
Atlanta, GA 30339-0049
(404)933-5539

Ron Norris
316 Cox Road

Newark, DE 19711
(302)774-3504
(302)999-3969 FAX

Gregory B. Teague, Ph.D.
New Hampshire-Dartmouth Psychiatric Research Center
2 Whipple Place
Lebanon, NH 03766
(603)448-0126
(603)448-0129 FAX

Robin Turpin, Ph.D.
JCAHO
1 Renaissance Blvd.
Oakbrook Terrace, IL 60181
(708)916-5923
(708)916-5644 FAX

Jonas Waizer, Ph.D., Chair
Senior Vice President
F.E.G.S.
510 Sixth Ave.
New York, N.Y. 10011
(212)366-8024
(212)366-8033 FAX

EXECUTIVE SUMMARY

In its Seventeen year history, the Mental Health Statistics Improvement Program (MHSIP) has enjoyed advancements in two areas: enhancing mental health statistics and information Systems and supporting the use of statistical information in the management and study of mental health programs. Data-based decision making will probably not continue to advance, however, unless there is a significant increase in use of the data generated by MHSIP-consistent systems.

Input from the MHSIP community has suggested that performance indicators (PIs), derived from the content of MHSIP can provide managers with important information and analytic capability. PIs can also reinforce accountability, evaluation and data-based management decision making.

At the request of the MHSIP Ad Hoc Advisory Group, a task force was convened and charged with

o the development of a conceptual framework and a model of PIs that can be derived from the MHSIP content, and

o preparation of a report that incorporates the conceptual framework, choice of indicators and their use, and recommended format for presentation of indicator data.

The Task Force consisted of a broad representation of potential users of PIs, including SMHAs, public and private service providers, families of and advocates for persons with mental disorders, academia, and the Joint Commission on Accreditation of Health Organizations (JCAHO).

Early in its deliberation, the group decided to

o emphasize the multiple perspectives and differential needs for indicators,

o underscore the importance of going through the process of participatory development of a System of PIs, and

o present a set of scenarios that reflect multiple perspectives and corresponding differential sets of indicators.

Purpose

The purpose of PIs is to communicate meaningful, important, data-based performance information in concise terms. The information communicated as ratios or rates can reflect available resources, processes, or outcomes, in order to assess general performance, assist and support management functions, and maximize responsiveness to service needs or legislative mandates.

PIs and values. PIs should reflect the performance of the mental health service system in the areas that are most valued by the different constituencies who feel ownership or have stakes in the service System.

PIs and policy. Policy is typically an articulation of what ought to be and of what is or is not desirable. In this sense, policy codifies the values in human services by guiding the actions of the service delivery System. Since PIs, by definition, reflect actions, they should also reflect policy implementation and its effects.

PIs and responsiveness to the needs of people with mental disorders. The mandate to the mental health system is to meet service needs of defined populations and subpopulations. PIs help determine how persons with mental disorders, their families and their communities are being served.

PIs and resources. Every translation of a policy into action requires the use of resources, whether money, physical property, or staff time. One of the most fruitful uses of PIs is to measure the volume and efficiency of the resources consumed.

PIs and impact. PIs are invaluable mechanisms for ascertaining whether policies and use of resources have produced the desired effects and whether the effects were of sufficient benefit.

PIs and decision support. PIs are powerful tools for decision support. They are a robust and parsimonious way of reducing and presenting a large volume of data in a way that assists in making decisions about, for example, allocation of and accountability for resources, compliance with mandates, choice of service providers, etc.

Effective PI systems. As PIs are developed, it is useful if the following principles are forefront.

o Explicit relationship to values and policy

o Sensitivity to context and environment

o Participatory development

Incorporating these principles into ongoing Systems is likely to increase their effectiveness.

A Conceptual Framework for Performance Indicators

There are a few key *principles* that should govern P1 systems.

o PIs should reflect values and policies.

o The process of selecting PIs should ensure flexibility and the ability to shift to measures that reflect the most important issues and policies at a given time.

o The logical sequence and process of selecting PIs consists of four phases:

oo *Identification* of the "need to know," which is determined by policies, objectives and political dynamics;

oo *Incorporation* of information needs and a conceptual framework for PIs;

oo *Articulation* of management and stakeholder questions and concerns; and

oo *Development* of a corresponding set of agreed-upon PIs to assess policy implementation and answer the most important questions.

o PIs should be ratios and rates-not raw numbers-and should not be merely descriptive statistics; rather PIs should help show the degree to which service Systems perform as intended.

o PIs are largely organization-based because organizations are the source of the data. This should not diminish the importance of consumer-centered Systems of care. Being organization-based, PIs provide Systems managers with leverage for shaping the performance of individual organizations.

o Frequently, PIs raise questions, e.g., about causes of the performance level shown. A combination of related indicators could suggest an explanation for the revealed performance level.

o Performance is multi-faceted, and different aspects of performance are not independent of each other. Because of possible trade-offs, an emphasis on only one aspect of performance (e.g., efficiency) could be at the expense of another aspect (e.g., effectiveness) and, therefore, should be avoided.

There are several types of *deteminants* of Systems of PIs:

o The selection of specific PIs is influenced by the *perspective* of the reviewing body, i.e., the entity conducting the analysis.

o The three primary *uses* of PI data are 1) shaping of behavior, 2) gauging performance in terms of congruence with standards or contractual agreements, and 3) analytic, as in research.

o Performance is always measured from one of three basic comparison points:

oo Performance of the same unit over time,

oo Performance level across units, or

oo Comparisons against an a priori value, such as a goal, standard, or norm.

Concerns to be addressed via PI data could center around broad subjects of analysis, such as target populations, services being provided, quality of services provided and the viability of organizations providing the services.

Proposed Paradigm

Mental health system managers have concerns that reflect three *dimensions of performance* - responsiveness to need for services, efficiency, and effectiveness - and typically focus their concerns on two levels of measurement, or *units of analysis* - client characteristics and behavior, and organizational characteristics and behavior. Integrating the type-of-performance with level-of-concern results in a two dimensional matrix of performance indicators which compare across either client type or across organizations.

Responsiveness is the congruence of the service structure, activities, and clientele with assessed needs; efficiency is the volume of output achieved, given the resources provided; effectiveness is the extent to which the outcomes were achieved through use of the available resources. Responsiveness, efficiency, and effectiveness are assessed across client cohorts, or other groups of recipients of service, or across organizations and organizational units.

System integration is another concern within each cell of the three-by-two matrix of dimensions of performance by units of analysis. System integration includes clients' needs for generic services, linkage, and referral patterns and other measures of the degree to which performance has transcended a traditional organizationally-based system of care.

The matrix purposefully does not include measures of compliance. This exclusion is based on the notion that any measure of performance can be built into legal requirements, policies and procedures, and standards of care. Compliance is, therefore, not a dimension of performance and its measures could reflect performance in any cell of the six~eli matrix.

In sum, the proposed paradigm consists of a matrix of three dimensions of performance by two levels, or units of analysis. System integration is another level within each cell of the matrix. Compliance could involve any of the cells and subcells of the model.

Development and Utilization of Performance Indicators

Three themes permeate the topic of development and utilization of PIs:

o the need to focus on utilization and its relations to and impact on policy,

o the importance of the process of the design and implementation of PI systems,

o the logical and most productive sequence of development of a PI System. The wide range of intended uses of PI data tend to fall into three categories:

oo continuous improvement of performance and service delivery through both comparison of one's performance with that of others and through periodic assessment and a self-correcting process,

oo gauging performance and service delivery against contractual agreements, regulations, or standards -- to maintain or secure accreditation, licensure, or future contracts, or

oo participation in basic or applied research for describing, predicting, or explaining performance.

The motivation to develop and use PIs appears to be either theory-driven (i.e., how things work best), or policy-driven (i.e., how things ought to work). Becoming fully aware of these motivations and articulating the intent behind the development and implementation of a P1 System are essential to maximum utilization of the resulting data.

The ideal environment for the development of a system of PIs is one in which:

o Intents of all stakeholders are articulated.

o There is a culture of respect for and constructive use of data.

o Changes are accomplished through participatory development.

o Resistance is reduced through an open discussion of any misgivings about the PIs and implementation of safeguards that address those misgivings.

The development of a sound System is contingent on a shared process tailored to the specific context and situation. A tailored, participatory process integrates the interests of all stakeholders. It also maximizes the relevance of the formal set of indicators to whatever is most important in a particular situation at a particular time.

A sound and logical sequence of developing a PI System consists of four components:

o identification of the developers of the system and their perspectives;

o articulation of stakeholders' questions, issues, and concerns to be addressed by the PIs;

o selection of PIs; and

o advanced decisions about the use of resulting data.

Different audiences may require different presentation strategies. The presentation of PI data should be designed to facilitate the reading and communicate the meaning of the results. Often, this is best accomplished through graphical portrayal of the findings.

Carefully crafted decision rules should be developed in advance and applied uniformly in the utilization of PI findings. The rules should be used to define high and low performance and determine how levels of performance should be related to consequences, such as the allocation of resources and service contracts.

The final step in the implementation and use of PIs is the evaluation and possible revision of the PI system.

Practical and Technical issues

A host of practical and technical issues should be considered in the development and implementation of a PI system. Addressing the following issues can help avoid common pitfalls and enhance the system to be implemented.

o *Planning* - political and organizational considerations related to the role of quantitative measurement in a given management system and how to go about selecting and implementing PI's

o *cost. burden and system capacity* - direct and indirect costs of the development, maintenance and management of both on-going and new data collection that is necessary to produce PIs

o the need for *multiple indicators* that reflect all important aspects of performance with minimum duplication and redundancy

o *technical measurement* issues related to error, validity, reliability, reactivity, range, variation and sensitivity to change, sensitivity and specificity of classification measures and appropriate interpretation of results

Illustrative Scenarios

Seven scenarios of PI use are identified to underscore the themes presented throughout the report, and to illustrate the logic of relating perspectives, management questions, corresponding PIs, and the use of resulting data. There are no "right answers" or "right indicators" for each of the scenarios. The presented sets are examples of potentially meaningful and useful

measures.

The sample scenarios reflect a wide variety of organizational structures, perspectives, concerns, stakeholders, and policy positions. The seven scenarios are:

o an SMHA;

o a local area, such as a county that contracts with private service agencies for the provision of mental health services;

o the perspective and concerns of both consumers and their advocates;

o a psycho-social rehabilitation service agency;

o a private, multiple-location mental health service corporation;

o a private managed care organization; and

o assessment of compliance with two legislative requirements of PL99-660.

Congruence with MHSIP Data Standards

The process presented in this report is driven by MHSIP, but is likely to result in an extensive set of useful indicators, some of which draw upon data that are outside the current MHSIP content. Data not currently in MHSIP, but of interest to mental health organizations in generating PIs include the following:

o need data that require general population statistics and epidemiological findings;

o support and generic services statistics, consisting of data collected by other health and human services entities that are involved in providing services to persons with mental disorders;

o consumer outcome measures, such as increased level of functioning, improved quality of life, and satisfaction with services received

Desirable systems of PIs will, therefore, have three relevant features:

o Their content will include data reflecting services and service organizations beyond those operated and funded by SMHAs.

o The measures will reflect current trends to include information on all services to consumers of the specialty mental health system, mental health services provided to consumers of other health and human service agencies, information that fosters decision support, and information that promotes a consumer-centered service system.

o As much as possible, data items will be consistent with the MHSIP standards and will enable cross-system comparisons.

INTRODUCTION

Background

In its fifteen year history, the Mental Health Statistics Improvement Program (MHSIP) has enjoyed advancements in two areas: enhancing mental health statistics and information systems, and supporting the use of statistical information in the management and study of mental health programs. Three factors bolstered MHSIP's recent accomplishments:

o endorsement of MHSIP standards by the National Association of State Mental Health Program Directors (NASMHPD)

o documentation of both the philosophy and standards of MHSIP in the National Institute of Mental Health (NIMH) publication *Data Standards for Mental Health Decision Support Systems* (FN-10)

o funding by NIMH of MHSIP implementation grants to the states

Data-based decision making will probably not continue to advance, however, unless there is a significant *increase* in the use of data generated by MHSIP-consistent systems.

Members of the 1990 MHSIP Implementation Task Force and many in the MHSIP community have suggested that PIs, derived from the content recommendations of MHSIP, can provide managers with important information and analytic capability. PIs can also reinforce accountability, evaluation, and management decision making. The group that produced the fiscal indicators in FN-10 took a preliminary step in this direction. A similar, but more comprehensive approach could offer more, showing how individual items of content can be combined into a variety of ratios, indices, and formulas to measure different aspects of performance.

Most state mental health agencies (SMHAs) have made progress implementing MHSIP. Continued support for implementation, however, will require the demonstration of management payoff, that is, assurance that MHSIP content improves acquisition, distribution and defense of resources, and the monitoring and evaluation of service programs. PIs derived from MHSIP-consistent statistics enable managers to pose and analyze complex questions and realize measurable benefits. To that end, the MHSIP Ad Hoc Advisory Group recommended the convening of a task force consisting of representatives of SMHAs, public and private service providers, families of and advocates for persons with severe mental disorders PSMD[1], academia and the Joint Commission on Accreditation of Health Organizations (JCAHO). In response, Jack Burke, M.D., Director of DASR, NIMH[(1)], approved in March 1991 the convening of and funds for the Task Force on the Design of Performance Indicators Derived from MHSIP Content. The Task Force consisted of the eleven members who prepared this report, whose names and affiliations are listed on page *i*.

Approach

The overall goal of the Task Force was to design, develop and deliver a document that shows how the content of FN-10 can be used to generate a variety of PIs for mental health program management and decision making. The charge to the group included the following:

o Review the charge and related documents, develop a plan and propose a model for a set of mental health PIs.

Address the purpose of PIs, deciding whether the indicators should be descriptive or valuative, whether the model should be multidimensional and, if so, what dimensions should be

represented, what additional data (other than MHSIP) to incorporate, etc.

o Develop a consensus concerning the number and choice of a minimum, core set of indicators per category (component) of the model and suggest formats for the presentation of indicator data.

o Prepare a report for the MHSIP Ad Hoc Advisory Group that incorporates the conceptual framework of the PIs, choice of indicators and their use, recommended format for presentation of indicator data, and recommended support activities and processes.

As directed by the MHSIP Ad Hoc Advisory Group, the Task Force approached its assignment in several phases. First the group reviewed its charge, work plan, expected schedule, tasks, assignments, and general orientation. Next the group reviewed and discussed all relevant background material, including the following:

o results of the survey of technical assistance needs

o pertinent developments in provider and auxiliary level settings

o a sample of models of PIs

o other prior work in the area

The most pressing questions, the answers to which would determine the final product, were identified as follows:

o Who will be the audience for PIs to be developed?

o What is the desired model?

o What specific data should be included?

The group's activities, individually and in five meetings over fifteen months, consisted of the discussion and development of the answers to these questions and of alternative models. The models varied in complexity (e.g., number of dimensions), content (specific ratios), and emphasis (e.g., whether to include compliance with mandates). Deliberations also weighed the merits of a general model against those of models tailored to specific situations, as exemplified by The State Comprehensive Mental Health Services Plan Act of 1986 PL99-660).

Early in its deliberation, the group re-examined its charge and intended product. majority of the group advocated, one, an emphasis on multiple perspectives and differential needs for indicators and, two, a product consisting of scenarios of such perspectives and corresponding differential sets of indicators. This decision and the rationale for it are explained later in this document.

Literature Review

The task force reviewed the available literature on the development and use of performance assessment systems and PIs for mental health services. while PI models have been used in many areas (e.g., economics, engineering, agriculture), the review for this report is limited to mental health services.

Overall, the literature on PI models in mental health reveals a peak of studies in the early 1980s (Hadley, et al., 1983, Kimmel, 1983) and a renewed interest in the late 1980s and early 1990s (Skinner, et al., 1988, Anderson, 1991, Barrett, et al., 1992). An emerging literature on the use of PIs for Total Quality Management (TQM) has added to emphasis on productivity and efficiency with a focus on consumers' values, employees' contributions, and quality as an operating strategy (Deming 1986, Walton 1986, Peters 1987).

Windle (1986) defined program performance measures as "operational specification of how well an organization is functioning along one or more dimensions that represent agreed upon goals or values of the program." He added that "these measures are expected to be quantitative, objective, and calibrated against some standard(s) that permit comparison within organizations over time and between organizations participating in the program."

What do managers and other stakeholders need to know about the performance of mental health service programs? Several systems of PIs described in the literature present different answers to this question. Jacobs and Thompson (1986) described the NIMH's Operations Management System (OMS) for community mental health centers. In the OMS system, PIs were selected as indicators of three goals: service accessibility, service organization's financial viability, and productivity/efficiency. Sorenson et al. (1986) have developed a set of indicators that addresses four areas of concern: revenues, consumers, staff, and services. Kamis-Gould (1987a, 1987b) described the New Jersey PI system, which incorporated Sorenson's areas of concern as the databases and as the sources for performance measures. The emphasis in the design of the New Jersey system was on what the measures were to indicate, i.e., the dimensions of program appropriateness, adequacy, efficiency, and effectiveness.

The literature identifies multiple uses of PIs. Hadley, et al. (1983) described a system of PIs that was implemented in Pennsylvania and used in allocation of state funds to county mental health programs. The intent of that system of PIs was to reinforce the reduction of the number of admissions to state hospitals, increase efficiency, expand the range of services, especially to PSMD, and encourage prompt submission of reports.

Barrett, Berger, and Bradley (1992) described the Colorado model of performance assessment and its implementation. That system consisted of five dimensions of performance: financial viability, productivity/efficiency, community responsiveness, comprehensiveness, and consumer/patient outcomes. The Colorado system was designed for performance contracting. Recognition of potential pitfalls lead to a stage-wise implementation and the use of safeguards, such as the development of baseline data and avoidance of sanctions for a predetermined period.

Anderson (1991) described PIs as a shared "vocabulary of performance" for internal management and continuous quality improvement (CQI), benchmarking via comparisons with other service providers and accountability to consumers. He emphasized that the choice of PIs should be determined by the mission of the organization and the suitability of each measure to gauge whether the organization carries out its own mission.

Rosen, Miller and Parker (1989) promoted the use of PIs for the development of standards of care. Leff and Natkins (1985) described models of funds allocation, where indicators reflected performance in terms of equity and need for service. Russell and Cole (1987) promoted the need to assess outcomes, impacts, and effectiveness.

Kamis-Gould (1987a) stressed two issues: one, the importance of *the process of* the design, development, and implementation of PIs and, two, likely tradeoffs between some facets of performance and others. In the system she described, decision rules about desirable and undesirable performance always involved more than one dimension of performance, more than one reporting period, and a confidence level of two standard deviations. Wholey and Hatry (1992) also suggested the need for the development of recommendations on the process for

implementing effective systems of performance, monitoring, and assessment.

Kimmel (1983) as well as Wholey and Hatry (1992) stressed that there is no "right set" of PIs, that performance measurement and monitoring in mental health focuses attention on some behaviors and outputs (and not on others), and that multiple factors contributed to the selection process. Kimmel saw the selective attention and tailored Systems as a weakness and a factor in the dynamics where agencies try to "game" the system and distort data to appear favorably. Wholey and Hatry suggested that "creaming" (serving "nice consumers and those who are more likely to improve) and "gaming" could be minimized by the creation of realistic expectations, participatory development of P's, implementation of a balanced system of PIs, and using PIs for comparisons of only comparable programs and consumers. Skinner et al. (1988), described another risk of using PIs by documenting amplification of errors through the use of indicators. The authors acknowledged, however, that audits could alleviate this problem.

What can be learned about PIs from either the literature, or from the collective experience of members of the task force? Most likely, PIs are here to stay. The quest for performance assessment and accountability to all constituents is gaining momentum, as exemplified by federal legislation, e.g., The State Comprehensive Mental Health Services Plan Act of 1986 (PL99-660) and Senator Roth's bill "Federal Program Performance Standards and Goals Act of 1991" (S. 20), and by various state initiatives. National agencies insist on monitoring and evaluating the public use of federal funds and compliance with legislative requirements, while consumers and their advocates make their own assessments.

Some systems of PIs, e.g., the one in New Jersey, were implemented and then discontinued. However, both ongoing and discontinued systems, e.g., the one in Pennsylvania, have shown multiple uses of PIs and their success as change agents and as instruments for shaping the performance of mental health service programs. Evidently, all management, whether short- or long term, consumer-focused, value-added, or market-driven, needs accurate and timely information to communicate performance, stimulate improvement, increase confidence, gain and justify resources.

There are several lessons to be learned from past experiences.

1. PIs are important as operational measures of key policies, policy implementation, and successive approximations toward the desired system of care, within a particular context.

2. PIs should always be dynamic, valuative ratios to be used in cross sectional or longitudinal comparisons. In other words, PIs should be measures of assessment, rather than description, with dear meanings of whether high, or low values are desirable. A corollary of this second point is the importance of longitudinal data and PIs and their usefulness in steering change and development.

3. Performance is multi-faceted and systems of PIs should include a balanced reflection of performance--neither only strengths, nor only weaknesses.

4. The process of choosing and implementing PIs is very important. The choice should be based on a consensus about how the system of care should be changed. Implementation should be through participatory development and collaboration among all stakeholders. This is especially important because mental health programs usually operate within a political system and because data, PIs, and information on performance are often exercised in struggles for influence.

5. PIs are powerful tools for policy implementation and shaping the behavior of service programs. To prevent misuse of PIs, reliability and validity of measures must be demonstrated, the burden of data collection should be minimized, and resulting information should be used judiciously.

This report departs from and adds to the existing literature in two ways: one, it offers a synthesis and (at least partial) resolution of issues raised in the literature, and two, it promotes a

generic model for the development of a sound system of PIs, including examples. The proposed model is generic and applicable to diverse situations. Its content, however, is likely to vary in response to different situations (context, policy issues, etc.), as illustrated in the examples. The development and implementation of systems of PIs that follow the proposed model should advance the implementation of MHSIP and foster data use in management and policy development.

About Performance Indicators

There are four underlying assumptions about PIs that guide this report.

First, PIs should be developed as a reflection of specific values and policy concerns of the developers. It follows that the optimal choice of PIs often changes as policies and knowledge bases change. Because P's are developed for managers to be able to monitor policy implementation, relevant specific measures are likely to be unique to the context and policy arena in which they are created. In most cases, PIs should only be viewed in the organizational and policy context from which they are derived, and interpreted contextually, not in isolation.

Second, PIs can be used by different audiences for a variety of purposes. They can be used to self-manage, to increase quality, or to improve productivity. PIs can be used by state, county, or private insurers to allocate and manage resources in light of the policies to which they subscribe. PIs can also be used as diagnostic tools in the evaluation of Systems or agencies.

Third, PIs are always comparative in nature. They provide for comparisons of similar organizations or consumer populations, comparisons of the same organization or consumer group over time, comparisons of organizations against requirements, goals set by the agency, or by an external environment. Thus, PIs are analytic and evaluative, rather than descriptive.

Finally, it must be kept in mind that PIs are inter-related and that one aspect of performance (e.g., efficiency) is not independent of others (e.g., effectiveness). The reading and interpretation of PIs should, therefore, be treated as a system of related measures and never in isolation.

In sum, the primary way to further the implementation of MHSIP and foster data-based decision-support systems is through the implementation and use of PIs. This report is the culmination of the work of the Task Force on the Design of Performance Indicators Derived from the MHSIP Content, convened by the MHSIP Ad Hoc Advisory Group and supported by NIMH. This document offers a discussion of key issues related to the design, implementation, and use of PIs and promotes a generic, multi-dimensional model of PIs for mental health service systems.

Overview of the Report

Chapter II articulates the purpose of developing and operating a PI system. A conceptual model for such a system is presented in chapter III and detailed in the paradigm that is presented in chapter IV. Chapter V outlines everything involved in creating a PI system and its use, including presentation of results. This is followed by a discussion in chapter VI of practical and technical considerations. Chapter VII consists of seven scenarios of PIs that represent a broad range of perspectives, concerns, management questions, and PI measures. The last chapter is a brief discussion of the congruence of PIs and MHSIP data standards.

II. PURPOSE

The purpose of PIs is to communicate meaningful, important, data-based performance information in concise terms. The information communicated as ratios or rates can reflect processes (e.g., staff productivity), outcome (e.g., average functional improvement per consumer discharged), or resources (e.g., full time equivalent direct service staff per 100,000 population). The

most common goals of PIs are the following:

o Assess general performance.

o Assist and support management in allocating resources, monitoring services, and evaluating impacts.

o Account for and assess responsiveness to service needs or legislative mandates.

Monitoring and assessment functions can be performed by several groups:

o managers

o professional peers

o consumers

o advocates on behalf of consumers and their families

o quality assurance organizations

o funding authorities

PIs can be thought of as a funnel that transforms several sources and types of data into concise, useful assessment information.

Performance Indicators and Values

PIs should reflect the performance of the mental health service system hi the areas that are most valued by the different constituencies who feel ownership or have a stake in it. Some stakeholders may be interested in whether the system relieves human suffering; others may value whether the system turns a profit while it operates; still others may care whether it serves a reasonable number of people with the resources it has available; and most stakeholders will have a multiple set of concerns about the service system. Systems of PIs should consider this range of stakeholders, what they value about the operation of the mental health service system and what they want to know about it. The PIs should be as responsive as possible to these concerns.

The purpose of PIs is to reflect policy formation and implementation in three general dusters of values and concerns shared by most stakeholders:

o responsiveness to the needs of persons with mental disorders,

o the use of available resources, and

o the consequences and impacts of policy.

Performance Indicators and Policy

Policy is typically an articulation of what ought to be, what is or is not desirable. In this sense, policy codifies the values in human services by guiding the actions of the service delivery system. Since PIs, by definition, reflect actions, they are intimately related to policy. This relation is one of reciprocal influence. Policies define what is valued and thus what should be measured; PIs measure how well a policy is working. Both should stem from a single perspective. The examples at the end of this report illustrate policies that drive the generation of sets of PIs from different perspectives. Data gathered over time will confirm or question whether each policy is having a desirable or acceptable effect.

Sometimes, values outface technology and have no ready measures, information systems, or specific pieces of information that can generate a numerical indicator for an area that is valued. For example, consumers may value the extent to which they have been accorded fundamental respect and dignity by the service system. Although a value of genuine significance, neither the mental health statistical community, nor consumer groups are yet able to articulate a formula or algorithm to measure it. In 1992, MHSIP began to collaborate with consumers to address this problem and begin to internalize consumer-based perspectives into the data standards, which are a fundamental component of MHSIP. The paradigm proposed in this report will be able to incorporate such measures, once available.

Performance Indicators and Responsiveness to the Needs of PSMD

PIs help determine how well PSMD, their families, and their communities are being served. while policies are needed to structure and direct the activities of mental health services (e.g., establish priority consumer groups to which to target specific services), the "bottom line" is the people. PIs are a practical and valuable tool for addressing the bottom line. They deal ultimately with the people by ascertaining whether:

o the service system can adequately meet their needs

o the right persons have been reached by the system

o the priority consumers are served

o services are appropriately tailored to the needs of consumers.

Performance Indicators and Resources

As policy is shaped, articulated, and operationalized, it takes the form of regulations, procedures, statements of goals and objectives, the establishment of priorities, governance and advisory structures, and prescriptions about program behaviors. virtually every translation of a policy into action requires the use of a resource, e.g., money, physical property, or staff time.

One of the most fruitful applications of PIs is to reflect the volume and efficiency in the use of these resources. PIs become an empirical shorthand that accounts for how well resources

were used in translating policy into action.

Performance Indicators and Impacts

In nearly every instance in which policy or resource consumption is considered, there is an accompanying concern about impact. That is, did the policy or the use of resources produce the desired effects? Was the effect of sufficient benefit? PIs are invaluable mechanisms for answering these questions and handling a complex array of considerations. The policy (an operationalized value statement) sets the stage for what is supposed to happen and thus suggests the types of indicators that should be examined. The indicators show whether appropriate services were available and the volume of resources that was consumed. By adding an outcome, impact, or quality expectation, a stakeholder gains some insight about four impacts:

o whether the policy is sufficient, reasonable, and is being implemented

o the resources consumed to make it happen

o whether the effect was intended and of acceptable magnitude

o whether the service recipients have benefited as intended

Performance Indicators and Decision Support

It should be apparent from the preceding section that PIs can reflect a great deal about the operation of a particular service program and the service system. Perhaps the best way to summarize this is to say that PIs are powerful tools for decision support. They are a robust and parsimonious way of reducing and presenting a large volume of data in a way that assists in making decisions.

As emphasized above, PIs need to be grounded in values and policies. PIs will help decision makers feel comfortable with a policy and realize what is not working and what deserves to be reconsidered or managed better. They can even expose an area in which no policy has been articulated.

One major responsibility of decision makers is to assure that the service system does what it is designed to do, e.g., whether it is responsive to the needs of PSMD and of other consumers. PIs enable managers to monitor and evaluate performance in this domain and make decisions accordingly.

Another key responsibility for decision makers is the allocation and monitoring of resources. PIs shed light on the use of resources and, through periodic accounting for the consumption of resources, become a *de facto* monitoring tool.

One of the most telling pieces of information for decision making is whether the impacts produced by operations are ones that are desired and of acceptable quality. if so, the decision maker has a knowledge base for repeating the impact and possibly improving It. if not, the PI prompts further questions, examination of other data, etc., to determine what needs to be done differently.

Finally, PIs should not be viewed as ad *hoc* or unique to each service program. If the data items are grounded in MHSIP or other data that are comparable across other providers, the

decision maker can compare performance not only within his or her own program, but with other similar programs. Such comparisons often shake the complacency of a manager who realizes that the program's performance looks much different when compared to others. On a more systemic level, the comparisons can provide the best holistic views of the system and its components. They can reveal, for example, that many providers are having similar problems, that some providers are efficiently producing high quality outcomes, or that some programs need intervention to preclude embarrassment or legal actions.

This iterative process of empirical measures, policies, and management activities is a uniquely connected system -- a nexus - in mental health services. If done well, everyone benefits, resulting in the ability to develop sound policies, select useful indicators, and use indicators effectively.

Effective Performance Indicator Systems

As PIs are developed, it is useful if the following principles are forefront.

1. Explicit relationship to values and policy

As discussed above, PIs will be of greatest credibility and usefulness if they are developed in a manner that is sensitive to the values of the stakeholders who will use them and if they can be explicitly related back to policy. The process by which this occurs is the focus of the remaining principles for effective PI systems.

2. Sensitivity to context and environment

The mental health field, as a whole, now has more automated information Systems in place and more data in use than ever before. This is manifested by recent activity around legislative issues such as comprehensive community mental health planning PL99-660 and 102-321) and the entire MHSIP effort.

Over the past decade, stakeholders in this field have become increasingly proactive in using data to advocate for and oversee mental health services. Examples include surveys used by the National Alliance for the Mentally Ill (NAMI) to advocate increased research and improved services, use of data by providers to improve clinical practices, and use of data by managers to ensure better accountability and justify budgets. Two fundamental insights unite all these efforts:

First, there is a concern to serve better the PSMD. Three factors have combined to focus attention on interventions, treatment, and supports that are less restrictive and more affirmative of human potential:

o a major shift away from inpatient care,

o dissatisfaction with what was accomplished in the twenty-year community mental health center era,

o a pervasive belief in the potential of psychosocial rehabilitation for persons with mental illnesses.

Second, resources are scarce and accountability for their allocation is stressed. Closer scrutiny is paid to resource use and efficiency, of course, but there is also a greater openness to a wider variety of constituent perspectives, of family members and consumers, participating in decisions.

PI systems that key into these contextual concerns and local circumstances will determine the developmental agenda. For example, the first of the above emphasizes indicators that track inpatient use. A relevant set of indicators would include the following:

o reduced hospital stay

o increased portal-of-entry controls/reviews

o reduced instances of recidivism

o inpatient care that is limited to only the most severe, dysfunctional and grave psychiatric disabilities

3. Approach and process

No one "right set" of PIs covers all situations or satisfies all perspectives. Effective selection and use of PIs depends on multiple perspectives of key stakeholders. In addition, the development of PI systems should accommodate the evolving nature of policy, the sometimes precipitous alteration of priorities and developments that evolve from scientific knowledge.

As different stakeholders are engaged in the process of developing PI systems, two things will occur. The first, the manifest function, is that their perspectives will be solicited. The second, the latent function, is that an education objective will be attained. They will come to understand and be willing to use the indicators. Furthermore, this understanding will cover not just those to which they have had input; each stakeholder group will come to appreciate aspects of the inputs and concerns of other groups. The achievement of both the manifest and latent functions will result in highly effective use of the resulting system.

Another gain that accrues from a collaborative approach to PI system development comes from the literature on total quality management (TQM). Systems that are used improve, self-correct, acquire credibility and develop advocates. With a broad array of users, there will be continuous monitoring of the results and of the PIs. Such continuous use of the system leads to sharpened policies, greater efficiency, and improved quality/impact at the level of service delivery.

Finally, although the input of many stakeholders has been emphasized, the ultimate burden of producing the data falls to the providers of mental health services. They should always be involved in the process of PI system development. Aside from the benefits of their input and the sense of ownership they will develop from the process, a data quality goal will be facilitated. Specifically, as providers participate and develop feelings of partnership, they will place a high value on generating data that are of sound quality to the Pi system and come to recognize that the data they release to the system are treated with confidentiality (when appropriate), with professionalism and for the accomplishment of system benefits. In short, they will see a salutary value rather than punitive burden.

Thus, the purpose of PIs is to integrate extensive and detailed data and to transform them into meaningful and useful information. The resulting information should assist all stakeholders in assessing whether policies and mandates have been implemented and whether they had the intended impacts. The most frequent assessments will be concerned with responsiveness to service needs, use of resources, and the impacts of services. The findings from these assessments should shape future policies and management activities. The next chapter presents a conceptual framework for PIs and leads to a discussion of a generic model, its implementation, and its use.

III. A CONCEPTUAL FRAMEWORK FOR PERFORMANCE INDICATORS

PIs are the vehicle for capturing and reflecting important characteristics and "vital signs" of mental health service delivery in a minimal amount of data. As such, PIs can be a portentous and useful management tool. Because PIs are potentially powerful, it is important to recognize and describe their key features. This chapter lists and expounds upon key principles concerning PIs and identifies several types of determinants that shape the design of a system of indicators. This chapter thus provides a conceptual framework for a generic model of indicators.

Principles

PIs concisely compare consumers, organizations, and their attributes. Embedded in the comparisons are implicit values of who should be receiving what services, when, where, at what costs, and with what effects and policies that are expected to bring about desired performance. The values may be implicit or explicit, shared or unique, and be varied in degree of subjectivity and loyalty to specific ideologies.

In using PIs, managers want to assure that their organizations do what they are supposed to do and do it well. Doing well includes both objective values, such as the desire for staff productivity and organizational efficiency, and more subjective, or ideologically-based propensities, such as preference for services in the least restrictive environment. The policies are likely to reflect the values of key decision makers and emerging issues that are important to their organizations. Since the importance of different issues and concerns shifts over time, the choice and employment of specific indicators are likely to change, as well. The process of selecting PIs should ensure flexibility and the ability to shift to PIs that reflect the most important issues and policies at a given time.

There is a connection and a logical sequence that ties shared values and the derivation of PIs. Values and principles determine priorities and the selection of goals to be achieved by the system of care. Policies and procedure are established to facilitate the attainment of the goals in a manner that is consistent with the values and principles. Therefore, similar values (in comparable environments) are likely to produce similar goals and policies, which in turn result in a similar set of selected PIs.

Often the sequence and process of selecting PIs consists of four phases:

1. identifying the "need to know", which is determined by policies, objectives, and political dynamics,

2. incorporating a conceptual framework for PIs,

3. formulating related management and stakeholders' questions and concerns, and

4. selecting and deploying an agreed upon set of corresponding PIs.

The relevant policies, in this approach, determine and influence the "need to know", the questions asked and the development of consensus about the indicators to be selected for the set. Findings produced by the PIs should be used in process and outcome evaluations and the results fed back into possible new management questions.

In all these cases, PIs should not be merely descriptive statistics, but rather help show the degree to which service systems perform as intended. For example, the number of elderly who

are consumers is a descriptive statistic; the proportion of elderly among consumers in relation to proportion of elderly in the general population is an indicator of proportional participation in the services offered and an indirect measure of service accessibility. Thus, PIs should facilitate change and serve as a proactive management tool for discerning whether wanted changes have taken place and whether unwanted phenomena have been eradicated.

PIs, whether measures that reflect a policy of meeting 'minimum standards or of maintaining market share, should be developed and implemented in response to the most important concerns and policies. At times, measures of the implementation of key policies can override the values and belief of individual managers. Unlike quality measures (e.g., low infection rate in inpatient setting, timely and comprehensive treatment plan, etc.), there is no universal agreement about value systems. Values change and new ones emerge, which is the reason for the importance of the process by which PIs are selected. Some of the differences in the selection of PIs can be attributed, for example, to the values of politically- vs. outcome-driven systems.

The ideal process of developing PIs requires that management make policies explicit, so that desired policy implementation can be operationalized and measured. This, however, may be problematic because some managers have difficulty formulating their own policies; others may prefer to keep their policies implicit, in order to maintain the flexibility to shift emphasis. While some managers may prefer to be presented with a set of rules (rather than go through the process of articulating their own policies), there is much value and merit to going through the complete process of articulating intent and translating it into sound operational measures, i.e., PIs.

PIs are largely organization-based. Organizations are the source of data and the resulting ratios are primarily a reflection of the organizations that capture and relay the data. This "reflection" can consist, for example, of both attributes of consumers served by the organization and of the treating staff. This focus on organizations should not diminish in any way the importance of consumer-centered systems. Measures that focus on organizations are important because they provide Systems managers with leverage for shaping the performance of individual organizations (e.g., through performance contracts) and, thus shaping the total service system.

PIs provide richness of information for multiple audiences and should be viewed within their relevant context. The context could be a matter of the audience's perspective, informational needs, etc., which would also determine the degree of required details. An analogy could be drawn to information presentation via maps, where a general map (e.g., of a state) is analogous to a total set of generic PIs, indicators produced on state or national level, and a detailed insert (e.g., of a city) is analogous to PIs that pertain to a tailored informational needs, e.g., about compliance with contractual commitments of a single organization.

PIs, whether directly or indirectly rooted in values and policies, always "indicate" or reflect a level (or degree). They should always be in terms of ratios and rates, not raw numbers. For example, staffing levels could be indicated in terms of direct service FTE, per 100,000 population, or per 100 consumers. It should not be raw numbers. Interestingly and frequently, PIs raise questions, e.g., concerning causes for the level shown; they rarely provide answers.

A combination of related indicators could suggest an explanation for the level revealed (and therefore guide decisions). For example, a measure of low proportion of minority consumers might be due to a low proportion of minorities in the service area. Services in a rural area, where clinicians have to travel a great deal to reach consumers, might explain a high cost per unit of service, etc. A desirable set of PIs can be used to teach people how to use them. The set should be small, but sufficient to reflect inter-related aspects of performance. if, for example, the cost per unit of service is high, the set should contain the measures that could be related and might explain the level of cost.

A particular type of decision, e.g., about allocation of funds, can be based on indicators reflecting different aspects of performance, whichever is to be reinforced. Thus, a state mental health agency (SMHA) may allocate funds based on high performance in compliance with a legislative mandate, in efficiency, in effectiveness, or a combination of the three. It is important, however, to keep in mind that facets of performance are not independent of each other and that, because of possible trade-offs, an emphasis on one aspect of performance (e.g.,

efficiency) could be at the expense of another aspect (e.g., effectiveness). Again, the ideal system should consist of a set of measures that is small enough to be meaningful and manageable, but reflect all essential and inter-related aspects of performance.

While performance standards are best developed via scientific data, client satisfaction must also be included. Measures of satisfaction are applicable to all scenarios of performance assessment, although they might originate in different intents. For example, a private corporation might monitor its quality and satisfaction standards in order to maintain its market share. The Joint Commission might monitor satisfaction because it views consumers' feedback as an essential outcome measure. The presentation of findings should be graphic and the interpretation thought provoking.

Determinants of Systems of Performance Indicators

PIs often derive from on-going information systems. They are anchored in specific points in time and are often used within the context of continuous monitoring. Examples of indicators range from simple ratios, e.g., unemployment rates, to complex ones, such as the Gross National Product. A similar range could be expected in mental health, where many aspects of performance are related to values, policy and compliance with mandates.

The selection of specific PIs is influenced by the *perspective* of the reviewing body, i.e., the entity conducting the analysis.

o Managers of a provider organization may review the performance of their own organizations, for many possible uses, using any type of comparison identified below.

o Administrators of a state mental health agency (SMHA) may assess the performance of mental health organizations in light of state regulations and funding agreements.

o Executives of a corporate entity may review the performance of their corporate agencies.

o External parties (e.g., JCAHO, NAMI, consumers, etc.) that transcend the management of an organization may have strong interest related to the service delivery of an organization, or a system of service, and use PIs to review providers' performance.

In the context of a decision support system, certain primary *uses* may be made of PIs. Three broad categories have been identified, while a number of subcategories and specific applications are possible within each of these.

o Shaping behavior/performance; revealing levels of performance, diagnosing possible reasons and identifying leverages for changing and improving performance. This use is also referred to as Continuous Quality Improvement (CQl).

o Gauging performance in terms of compliance with laws and regulations, or against either standards or contractual agreements. This use of PIs usually involves the application of sanctions, rewards, or punishments, as consequences of the degree to which performance conformed to expectations.

o Analytic, as in research, the purpose of which is to describe, predict and explain performance and contribute to generalizable knowledge.

Performance is always measured from one of three basic comparison points:

1. Changes may be assessed over time, e.g., comparison of the same unit of analysis from one year to the next.

2. Comparison may be across analytic units; consumer groups, organizations, organizational components, or a group of organizations, through the use of statistical, or normative base.

3. Performance may be assessed via comparisons with an a *priori* value, such as a goal, standard, or a norm (which in turn, can be based on best practice, average, or minimal acceptable level of performance).

As mentioned above, the specific PIs to be selected must be congruent with articulated values and policies and promote analyses in response to particular questions, issues, or concerns. The questions, or concerns, could center around certain broad subjects of analysis. For example,

o target populations - the recipients of services, e.g., whether, or not, intended consumers are being served, differential needs of consumer cohorts are being served, etc.

o services being provided - their appropriateness and effect

o quality of services provided

o viability of organizations providing the services

PIs appear in a matrix of *dimensions of performance* and *units of analysis.* The dimensions reflect mutually exclusive, major categories of what decision makers want to know, e.g., whether organizations serve the high priority populations, whether staff members are productive, whether costs are contained, whether consumers' level of functioning improve with services, etc. The units of analysis are the subject of the description and assessment, i.e., whether information compares consumer groups, or organizations. The content of the PIs, i.e., the types of data used in the numerators and denominators, will vary according to the combinations of and applicability to dimension of performance and unit of analysis. The next chapter delineates a generic model of PIs arrayed by dimension and unit of analysis. The content of the matrix cells varies according to specific contextual policies and issues, of which examples are provided in chapter VII.

IV. PROPOSED PARADIGM

Management Concerns

Managers use PIs to monitor the implementation of specific policies (concerns, goals, objectives, etc.) in those areas of the mental health system for which they have responsibility. It is critical that measures be developed specifically to provide information on concerns and policy agendas of a particular management entity or policy-making body. These concerns are translated into questions, and the questions are in turn operationalized into ratios of data, i.e., PIs. Despite the necessary specificity of an indicator to its policy or management context, there are general categories of indicators. Combining some of these categories in a conceptual framework of system responsibility and performance yields a tool for developing performance measures.

Mental health system managers typically focus concern on each of two levels of measurement, or units of analysis: consumer characteristics and behavior; organizational characteristics

and behavior. In addition, the interaction of consumers and organizations, as well as of both with their respective environments, are critical concerns. An important example of such interaction pertains to system integration, which is addressed below. For heuristic purposes, however, these other types of concern can be subsumed under a primary focus on either the consumer or organizational level. This triangle of concern can be visualized as in figure 1 (see next page):

The Paradigm

PI measures can also be grouped into three dimensions, or categories of performance; one, responsiveness to need for services, two, efficiency, and, three, effectiveness. In all cases such performance measures are expressed as ratios in order to permit comparison. They are ratios of such things as use, prevalence, resource consumption, or outcomes. Comparison is made across categories of consumer or of organization.

Integrating type of performance with level of concern results in a two dimensional matrix of PIs, i.e., of the possible comparisons across consumer type or across organizations. It is generally feasible to develop measures appropriate to each cell, but for any particular policy issue only certain cells may be relevant. In fact, designers of systems of PIs should not feel compelled to fill every cell of the matrix. The two-dimensional matrix is described in figure 2.

Figure 2. A Two-Dimensional Paradigm for Performance Indicators

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
Consumer Cohort			

Organization Cohort			

In this conceptualization, PIs are arrayed in terms of three dimensions of performance and two units of analysis. The content of specific PIs, the types of data used in the numerators and denominators, will vary according to the combinations of and applicability to dimensions of performance and units of analysis.

The three dimensions and two units of analysis are mutually exclusive categories that are defined as follows:

DIMENSIONS OF PERFORMANCE

1. *Responsiveness:* congruence of the service structure, activities and clientele with assessed needs. Assessments will often be based on the relevant service area population, but could be based on another specific set of individuals (e.g., PSMD, enrollees of a health maintenance organization, etc.) Relevant indicators in this dimension are ratios of an output measure (e.g., consumers served) over a measure of need (e.g., service area population, expected prevalence in the area, etc.).

2. *Efficiency:* the volume of output, or productivity achieved, given the resources provided. Indicators of efficiency are ratios of *an* input measure over an output measure. In analyzing *this* transformation of resources into output, indicators can be based on dollars, services, consumers, staff, or combinations thereof.

3. Effectiveness: the extent to which the outcomes, as they pertain to consumers, or groups of consumers, were achieved through use of the available resources. Indicators in this dimension are ratios of an output measure over an input measure. Such measures include an assessment of quality of life (QOL), level of functioning (LOF), clinical status, and feedback/satisfaction.

UNITS OF ANALYSIS

1. *Consumers:* consumer cohorts and/or other groups of recipients of service. Indicators pertaining to consumer cohorts can be used to examine sufficiency of services for a consumer population, that group PIs resource consumption and the effect of specific services on those consumers' QOL.

2. *Organizations:* sub-organizational units, program elements, organizational parts (e.g., human resources), or groups of organizations. Indicators may be used, for example, to compare the efficiency and effectiveness of two types of organizations, e.g., hospital-based vs. free standing.

For consumer measures of responsiveness, efficiency and effectiveness, 3 ratio that defines a type of consumer is always some numerator (e.g., persons served) over some denominator that defines a type of consumer (e.g., expected number of PSMD). For example, responsiveness measures could be in terms of the level of prevalence, access, incidence, or use by type of consumer, a comparison across types of consumers. Consumer types may be characterized on the basis of such variables as demographic characteristics, symptoms, service history, etc.,

but for consumer analysis, the denominator is always some type of consumer.

If one is interested in the responsiveness of one's system to PSMD, one could measure the percentage of PSMD of total consumers served. if one is interested in the efficiency of one's services to PSMD, it is possible to compare consumption of resources by PSMD consumers as compared to any counterparts. The same concept applies to outcomes, where one would measure change in level of functioning for PSMD vs. non-PSMD consumers. In every case, for the consumer row of the matrix, the analysis compares ratios where the denominator is types of consumers.

In a similar fashion, the development of organizational measures for responsiveness, efficiency and effectiveness requires the comparison of organizations (or organizational units) where the denominator is always one or more organizations. For example, in the responsiveness category, one might compare a percentage of PSMD served across several service agencies. Examining efficiency, one might compare the total resources consumed by PSMD as a proportion of total budget of each mental health center. Evaluating effectiveness, one might compare recidivism rates for PSMD across county administrative service units. Thus, assessment of organizations (or organizational aspects, such as staffing patterns) employs comparisons of ratios, where the denominator is a measure of organizations, a single organization, type of organization (e.g., hospital-based vs. free standing), or organizational feature (such as the composition of its staff).

Systems Integration

Within the six-cell matrix one could define other levels, or dimensions. They are not included in the formal structure, because a matrix with more than two dimensions is less intuitive and more unwieldy. Nevertheless, there is an additional focus of concern within both consumer and organizational levels that should be considered when developing indicators, i.e., the issue of system integration.

The dichotomy of consumer and organization has had much face validity for mental health system managers, but a simple two value model is insufficient. Managers oversee organizations, or components of organizations, and these entities, in turn, serve people who (particularly for PSMD) are served in the context of an extensive array of supports and other services. Also, other stakeholders are involved, including families, community members, and other organizations. Managers need to ensure continuity and quality of care within this more complex system.

Thus, the concept of system integration defines a further possible focus of concern within each cell. Needs, processes and outcomes relative to broader content areas can be evaluated at both consumer and organizational levels. For example, broader service provision, linkage, and referral patterns at the organizational level reflect consumer needs for housing, community integration and employment. PIs can be developed reflecting the degree to which performance has transcended a traditional, more clinical, intra-organizational interaction in response to a more comprehensively defined mission of the mental health system. The assessment of systems integration can be visualized as a third dimension of the matrix, or as a sub-division within each cell of the three-by-two matrix.

Assessment of Compliance

Clearly, one of the uses to which PIs have been put is to compare organizations to specified mandates, idealized goals or minimum standards. Thus, the issue of compliance and the performance of organizations against predetermined standard must he addressed. The matrix described, although quite useful for the development of "pure" dimensions of performance by units of analysis, purposefully does not include measures of compliance. The exclusion of compliance is based on the notion that any measure of performance can be built into legal requirements, policies and procedures, standards of care, and/or contractual agreements. This is true for need-based performance (e.g., the PL99-660 requirement to set need-based

quantitative targets of the number of PSMD to be served), for efficiency/cost containment (e.g., contractually agreed upon cost per unit of service), and for effectiveness (e.g., a state's agreement with a county to reduce the use of inpatient services in a state psychiatric hospital by a certain number).

Compliance, therefore, is not a dimension of performance and its measures could reflect performance in any of the six cells of the matrix and be applicable to consumers, organizations and/or Systems integration. Nevertheless, information about compliance is very important to managers on all levels of the system of care. Managers, when they create standards, are really creating ideals of system behavior. PIs compare actual organizations against these ideals. This can be as simple as a formula specifying that "half of all expenditures should go to PSMD," or as complex as the creation of an archetypal organization against which others are compared.

The proposed paradigm, therefore, consists of a matrix of three dimensions of performance, by two levels, or units of analysis. System integration is another level within each cell of the matrix. Compliance, whether congruence with legislative mandate, or consistency with contractual agreements, could involve any of the twelve cells and sub-cells of the model.

V. DEVELOPMENT AND UTILIZATION OF PERFORMANCE INDICATORS

Several state mental health authorities and national organizations have made a strong commitment to the development, deployment and use of PIs. This commitment was made public via articles in the professional literature, the National Association of Private Psychiatric Hospital's publication of "critical indicators" and JCAHO's promulgation of "Agenda for Change". To date, however, there is only scant research on the utilization and impact of PI systems and on the determinants of, or key to lasting successful systems. Evaluators, who are among the most vocal advocates of PI usage, have produced a number of theoretical manuscripts about utilization-oriented evaluation (Patton, 1986, 1988) and the importance of dissemination of findings (McLaughlin, Weber, Covert and Ingle, 1988) in any effort to foster utilization. Even this group, however, has had difficulty documenting the effects of various dissemination strategies on data usage. This chapter has three themes pertaining to the development and utilization of PIs:

o The need to focus on utilization and its relations to and impact on policy. Akin to Patton's "utilization-oriented evaluation," this theme underscores what PI systems are all about and that utilization and intended impact are the *raison d'être* for any data system.

o A message from the members of the Task Force to the reader about the importance of the process of the design and implementation of PI systems. While the utilization of PI results is the ultimate intent, all steps of the developmental process are essential for lasting implementation and successful utilization of results.

o The step-by-step delineation of a sound process of design and implementation of a PI system.

Motivation

Recently, a variety of organizations have expressed increasing interest in the use of performance indicator data. Although mental health services delivery organizations may typically envision internal uses for indicator data, state and local regulators, the Health Care Financing Administration, payers, the Joint Commission and others are all eager for access to reliable sources of performance data. There are both internal and external pressures to collect indicator information, because raw data alone are insufficient for a thorough understanding of consumer and organizational processes and outcomes. Performance indicators must be designed with a specific purpose or purposes for the data to be meaningfully translated into information and for supporting management decisions. The process of developing a system of PIs, maximizing its relevance and potential utility are described in the flow chart in Figure 3 (see next page).

Pis have a wide range of intended use, which tend to fall into three categories.

1. Continuous improvement of performance and service delivery, which includes two subcategories:

a. Assessment as part of a cybernetic process in which current performance is examined for feedback on consistency with the desired direction. Deviation from the desired direction triggers corrective actions, to be followed by another round of assessment, related correction, etc. This type of assessment can be conducted by internal, or external examiners. In either case, findings are used, without repercussions or penalties as informational mechanism for self-change, or continuous performance improvement. In this type of utilization, simple feedback becomes a change agent, as all PIs are ratios with known desired direction, that is, there is an agreement whether high is preferable to low (e.g., percent improvement) or vise versa (e.g., percent re-hospitalization or drop-out).

b. Comparison of one's own performance with those of others (rather than with a known, desired direction) to see whether change and improvement are indicated. In this case, managers compare their organization's performance with the performance of individual, similar organizations, or with normative data (average, best practice, etc.) derived from pooled data about similar organizations. Findings of performance below the norm create motivation for improvement.

2. Gauging performance and service delivery against a contractual agreement, regulations, or standards to secure accreditation, licensure, or future contracts. Both comparisons with the performance of other, similar organizations and with the same organization over time can be used to assessed improved status. Success can lead to either reward (e.g., renewed contracts) or lack of penalty (e.g., removal of government sanction). Failure may be drastic (e.g., loss of contractual agreement) or soft-edged (e.g., conditional accreditation).

3. Participation in basic or applied research for describing, predicting, or explaining performance. Each type of data, whether normative, inter-organizational, or longitudinal, can be utilized for research. Participation in research entails both short-term benefits (e.g., improved analytical capabilities) and long-term rewards of making contribution to public knowledge.

The intended use of captured and generated performance data is likely to impact motivation and incentives as organizations are more likely to produce data that could benefit them (e.g., increase revenues) and less likely to produce data that could be used for negative consequences. Participation in research is often related to either compensation for the burden of generating the data, or to personal interest of key decision makers.

The stimuli to develop and utilize performance indicators may originate from any source, but generally, motivation appears to be either theory driven (i.e., how things work best), or policy-driven (i.e., how things ought to work). Becoming fully aware of these motivations and articulating the intent behind the development and implementation of a PI system are essential to maximum utilization of resulting data. Since more often than not a system of care involves multiple stakeholders who would like multiple applications and uses of PI data, all these intents should be articulated and shared.

The motivation and investment in the development of PIs can be enhanced significantly if all stakeholders participate in all phases of the design and implementation. Participatory development maximizes the identification of what all constituencies want described and assessed and what consequences of assessment they fear; what safeguards to build into the process and how results should, or should not be used. For example, if service providers fear that revealed levels of performance could be used to penalize them financially, an agreement can be reached that PI data will be used for feedback only for a pre-determined period of time and that financial, or contractual, consequences will not be used until the conclusion of that "hold harmless" time. This way, participatory development could also maximize the buy-in by all parties and their sense of ownership in the resulting system.

The ideal environment, therefore, for the development of a system of PIs is one in which:

o Intents of all stakeholders are articulated and shared.

o There is a culture of respect for and constructive use of data.

o Changes are accomplished through participatory development.

o Resistance is reduced through disclosure of fears and implementation of safeguards that address those fears.

Importance of the Developmental Process

There is a good reason why this report does not provide a cookbook and short cuts for the selection of PIs, which is the conviction that the development of a sound system is contingent on a shared process tailored to the specific context and situation. There are many commonalities among different situations. These common aspects are best described in terms of dimensions of performance and units of analysis and are built into the generic model presented in Chapter IV. The tailored process that is promoted in this report fosters a shared experience by all parties and entails several features and related benefits.

o A tailored participatory process that encompasses the design and implementation of a system of PIs involves all stakeholders who articulate and share what they want to happen. It is conducive to resolution of differences and buy-in by all parties.

o A multi-step, sequential process that identifies and ties together values intents and selection of measures.

o A tailored developmental process that maximizes the relevancy of the indicators to what is most important in a particular situation at a particular point in time.

Sound Development of a Performance Indicator System

Once the decision has been made to utilize performance indicators, a plan should be developed for the: 1) strategic, tactical and operational design and implementation of the indicators, 2) management of the data, 3) assurance of integrity of data interpretation, and 4) utilization of results. Implementing the plan should result in the generic process depicted in Figure 3.

The process should be based on answers to the following questions:

1. Who are the developers of the system?

2. What are the questions, issues and concerns to be addressed by the system of PIs?

3. What are the PIs best suited to reflect the degree of implementation and impact of the most important policies?

4. How will the resulting data and information be used?

As described in Figure 3, the perspectives of the developers, the mandates with which they are charged and the values and policies shape the goals and objectives of the service system. Since most public mental health service agencies and systems function within a political system, the political environment and dynamics (e.g., downsizing of inpatient care, emphasis on consumer satisfaction, etc.) also enter into the process. The result of this component of the process is the identification of the initial "need to know

The second component of the process involves all stakeholders' questions, issues and concerns to be addressed by performance assessment. In this phase, managers and other stakeholders questions are arrayed into the generic model suggested in chapter IV.

The third component consists of the following:

o prioritization of the set of questions in each part of the model

o transformation of high priority questions into operationally defined variables

o creation of measures, in terms of ratios, that reflect levels of performance, determined by the set of management questions and concerns

o development of consensus among all stakeholders about the final set of indicators

o deployment of the indicators, i.e., performance assessment using the selected measures

The last component of the process consists of two parts:

1. handling of the data

2. application of the findings

The handling of resulting data should include careful interpretation of levels of performance, similarities and differences across organizations and over time, and presentation and dissemination of results. The application of the findings should close the loop, by feeding information back to the first component, by showing the degree to which performance reflects the implementation and impact of policies, and the degree to which the system has attained its own goals and objectives. In closing the loop, findings are likely to raise other stakeholders' questions that could be built into future performance assessment and shape future management.

Presentation and Interpretation of Performance Indicators

The presentation of PI data should be designed to facilitate the meaning of the findings and the interpretation of results. This is likely to be accomplished best through graphical portrayal of the findings. Five specific forms of graphical presentations are most appropriate for the presentation of PI data and their meaning.

Run chart: line graphs that relate the level of a PI in a time series. They are most suitable for the reflection of performance on the same variable over time.

Histogram: a graphical presentation of categorical data, such as the value of PIs of, for example, the same indicator across different organizations, or multiple measures of the same organization. It is particularly suitable for communicating similarities and differences, relative strengths and weaknesses and patterns within a distribution of data.

Pareto chart: a prioritized bar graph, most suitable for display of the order of prevalence of disorders, magnitude of different presenting problems, etc.

Control chart: a trend chart with statistically determined limits that predict bow much variation in the data is to be expected and when a reaction is warranted. Control charts are tools used to analyze and monitor process and outcomes. These charts graph trends in the data over time and include control limits (e.g., standard deviations) that delimit the extent of variation expected under normal conditions and indicate what would be considered a significant deviation. A control chart would be a good choice for displaying admission rates to a state hospital from the various catchment areas around it and two standard deviations above and below the mean of those rates, because it would identify rates that are statistically significant from the average and expected rate.

Scatter diagram: these five graphical tools are illustrated in figure 4 (see next page).

Figure 4. Graphical Tools

Use of graphs is increasingly advocated and practiced, but graphical presentation carries its own matters of technique, frequently ignored. Tufte (1983) provides a thorough analysis of the strengths and weaknesses of various graphical techniques in particular applications. Among numerous recommendations, Tufte cautions, for example, that small sets of numbers may be better presented as numbers. Instances of graphical excellence are almost always multivariate and give "to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space." Distances and dimensions on graphs should be proportional. Data should not be presented out of context - e.g., omitting the space representing low values from a graph. Extraneous decoration should be minimized, and unrelated dimensionality should be avoided altogether - e.g., three dimensional presentations of two-dimensional data. Although not all of Tufte's recommendations are pressing concerns for presenters of PI data, the goal of graphical presentations should be clear, efficient display, free of distraction or distortion.

Whatever the audience, presentation of data using techniques that are visually pleasing and helpful will enhance the utility of the information. Over all, the presentation of results should facilitate interpretation and provoke thought. This is because service systems are complex and performance levels could be interpreted in more than one way. For example, funds expended on staff, per 10,000 population, might reveal that more had been spent on white, than on minority staff. The dollar figures could represent insufficient minority human resources. They could also result from having newly recruited, younger minority staff who are paid lower wages than the "long timer" whites. Further examination of the data would be needed, in order to find the explanation of the differences. In all cases, to assure accurate interpretation and broad impacts, PI results must be disseminated to all stakeholders. Wrong interpretation of findings are also likely to be picked up and disproved if all constituencies review the results.

Dissemination of Findings

Much of what needs to be incorporated in presentation of the data is implicit in the foregoing discussion. If policies, concerns, data sources, algorithms, caveats, and results are presented or available along with interpretations or hypotheses, productive discussion can be encouraged and much unproductive reaction can be averted. If the processes of planning, development and implementation have incorporated the participation and investment of relevant stakeholder groups, and if the resulting indicator set is indeed linked to critical policy issues, the most important audiences will be ready-made. For particular PIs, such as indicators of client outcome, accompanying case illustrations, properly identified, can improve understanding of embedded concepts.

Different audiences may require different presentation strategies. Disseminators may want to think in terms of two key message types, each with its own objective (Joint Commission, 1992). *Knowledge communication* assumes that the consequences of actions taken by either individuals or the organization are unknown and that imparting factual evidence of these consequences would instill appropriate reaction and encourage appropriate action on either an individual or an organizational level. For example, an organization unaware that its indigent care rate was lower than would be expected might use this information to increase accessibility. *Persuasive communication* assumes that the audience is skeptical, and must be convinced to take action. Approaches to changing attitudes include sharing information about expectations or about linkages between behavior and positive or negative outcomes (Fishbein, Ajzen and McArdle, 1980). Effective persuasive communication is constructed to provide a set of arguments, appropriately matched with the belief system of the targeted audience, along with factual evidence designed to support the arguments.

A third type of impact on behavior and performance involves public disclosure of performance levels. Embarrassment from publicly-known low performance or pressure created by revealing deviation from the performance levels of other, similar entities creates strong motivation to shape future behavior in the desired direction. Thus, data and disclosure of performance information are powerful tools and can be used as leverage in shaping the behavior of service organizations. This step in the implementation of PI systems is one of the most important and, providing that previous tasks have been carefully executed, could serve as an aid to change agents, with relatively little investment of resources.

Decision Rules and Utilization of Findings

Carefully crafted decision rules should be developed in advance and applied uniformly in the application and utilization of PI findings. For example, in using a PI model that consists of three dimensions - responsiveness to need, efficiency and effectiveness -- an advanced agreement might stipulate the following decision rules:

o High and low performance on any one indicator is defined as two standard deviations above and below the mean, respectively.

o Designation of high performance is defined as high results (two standard deviations above the mean) on at least 2 indicators, in at least 2 of the 3 dimensions of performance, for at least 2 consecutive periods of assessment, and no low (two standard deviations below the mean) on any 2 indicators within any dimension.

In this example, an organization might have the absolutely highest indicator values in the area of efficiency, but not count as a high performer unless it is also high in either responsiveness, or effectiveness and does not perform poorly in any area. This kind of a rule is designed to underscore the inter-relationships among indicators and the need for reliable and stable findings. Allowing variability up to two standard deviations, this decision rule illustrates a somewhat less stringent approach to identifying high and low levels on any specific indicator. It is more stringent, however, in its designation of high performance. The example also illustrates that decision rules are a matter of agreement among all parties concerned. The rules could vary from one system to another and could change over time. The same system that permitted variability up to two standard deviations at the time of implementation, might tighten its rules after several rounds of assessment and limit variability to one standard deviation.

Ultimate Uses of Performance Indicator Findings

There are four categories of the application of PI findings.

o internal use as feedback in the process of continuous improvement

o use in services research in an attempt to understand the determinants of high performance

o input into modification of existing policies and development of new ones

o organizational and administrative sanctions

The fourth application represents major leverage in influencing and shaping the behavior of service-providing organizations. It could be in terms of new service contracts with a high performer, or discontinuation of such contracts with a poor performer. Applications could also be more thoughtful and complex. For example, a SMHA might want to integrate PI findings with data on assessed needs for mental health services in catchment areas served by the assessed organizations. In this case SMHA managers might decide that high performing agencies in high need areas will receive additional funds and will be contracted to provided expanded services. Low performing agencies in low need areas will be de-funded and their service contracts not renewed. Low performing agencies in high need areas will have to develop corrective action plans, and will be provided with technical assistance in order to help them improve. High performing agencies in low need areas might be recognized for their high performance and contracted to provide technical assistance to deserving low performing agencies. PI findings, in this example, provide managers with a major tool for both allocation of resources and the shaping of the service system.

Perhaps the ultimate test of the usefulness of a performance indicator system is its impact on the organization and its policies. Periodic review of the performance indicators within a policy framework can ensure the continuing relevance of individual indicators and of the PI system itself. Changes in the political environment, in specific policies, or in organizational performance or context may require the development of new or revised performance indicators as old ones achieve their ends, diminish in relevance, or even produce unintended, negative consequences. The principles and issues discussed in this chapter, critical to successful development and use of PIs, are equally important in the ongoing adaptation of the information system to the policy and service system environments. Continued sensitivity to these key practical and technical matters will help ensure that the PI system remains useful.

This chapter emphasized and detailed three topics: one, that the total design and implementation of a PI system should be oriented toward the utilization of resulting information, two, the importance of the process for a successful and lasting PI system, and, three, the components and sequence of the steps in the development and implementation of a PI system. Some details were also provided about the presentation, interpretation and utilization of results of any performance assessment data. Also mentioned were the utility of graphic display of data and agreement about decision rules in designation of high and low performance. The next chapter addresses important practical and technical considerations for a sound system of PIs.

VI. PRACTICAL AND TECHNICAL ISSUES

Previous chapters defined PIs, offering a structure for conceptualizing them. This chapter highlights several important practical and technical issues in planning and implementing PI systems as well as in using the results they produce. The development process contains many potential pitfalls at various steps. The savvy practitioner will be aware of and anticipate these pitfalls by heeding the caveats in this chapter.

Planning and Implementing

An important set of issues concerns the practice of planning and implementing PIs. These include both political and organizational considerations.

In the process of development of a system of performance indicators, there is a prior and usually tacit question: What is the proper role of quantitative measurement in a given management system? Although increasing credence and priority are given to objective demonstration of performance, top managers vary in the degree to which they are comfortable with

routine, widespread dissemination of data indicating organizational performance in potentially sensitive areas.

This report has repeatedly offered the view that effective indicators are policy-driven, and thus that they are factors in a political context. In the kinds of complex systems where PIs may be most relevant, information is an instrument of power in decisions about appropriate allocation of societal resources. Management style, strategic purpose, organizational context, and political environment all affect the way PIs can be introduced into a system, on the degree of penetration of a PI system into the organizational hierarchy, and on the type of content that might be measured.

Typically, as with implementation of TQM approaches, effective development of PI systems requires the joint commitment and collaboration of top management with a political vision and staff with a more technical, even technocratic orientation. Often the more technical subordinate will contribute some leadership in proposing and developing PIs and PI systems. This report emphasizes the critical linkage between PIs and policy, and cautions potential developers about the risk of basing systems too heavily on simplistic assumptions of rational planning and decision-making. A PI system must be adapted to its political context, notwithstanding the fact that this adaptation is necessarily mutual, since a successful PI system will also influence its context. From the point of view of a subordinate staff member with reporting system responsibility, for example, the political context includes both the management style and the political contingencies of the top manager.

The need for broad participation, discussed earlier, has much relevance here. Involvement of the appropriate range of stakeholders early in the planning stages provides not only opportunity to clarify attitudes and interests, i.e., relative to data on specific issues, but also to shape them toward a more productive consensus.

Cost, Burden, and System Capacity

Generation and use of data in a reporting system entails consumption of resources; part of the planning process is consideration of these anticipated costs in relation to anticipated benefits.

The use of PIs as a management strategy represents a significant commitment of system resources. System-wide introduction of new data collection to support PIs requires startup time for planning, training, and implementation, as well as additional ongoing effort. Even if a majority of PIs can be constructed from data generated for other purposes, such as financial accounting, program management, or clinical care, and if much of the cost is thus "buried," there may be significant effort in transmitting and analyzing the data, in communicating results, and in dealing with the effects of this information on the system. Much of this cost may be incurred from the PI set as a whole; the marginal cost of a single indicator is typically minimal. Nonetheless, the cost-effectiveness of each PI should ideally be estimated in advance, at least informally, and alternatives considered.

If a new indicator, however valid, requires substantial new training and data collection across a large organizational system, or, if a new indicator requires substantial programming or other data manipulation to achieve integration across the appropriate universe of organizations, implementers should evaluate the probable costs of these activities in relation to the probable benefits of the hypothesized changes to be brought about in the system through use of the data. In many instances, these costs will be worthwhile investments to achieve valuable ends; in other instances, less expensive indicators may be warranted.

Two other issues related to costs are redundancy and frequency. Developers considering inclusion of two measures of very similar phenomena need to weigh the advantages of mutual validation against the marginal cost of additional information gained; highly correlated measures may add little in analysis of results. Frequency of data collection should be relevant to the rate and importance of expected change; too short an interval provides redundant information, and too long an interval may prevent timely feedback for effective policy implementation.

Many costs will have tangible, financial value; others may take the form of intangible burden on staff or clients. A new evaluation or rating process, e.g., a measure of client need, eligibility or outcome, may not only take staff time and therefore represent consumption of a tangible resource, but the new activity may also impact staff morale. Such changes may have positive impact. If a measure provides improved focus or enhanced justification to clinical activity, it may improve morale along with services. Whenever possible, developers should pilot-test new data-collection methods to determine their effects in advance of full-scale implementation.

The issue of cost of measurement is increasingly salient as diverse groups become increasingly interested in questions of service outcome and system design and integration. Most current data systems still focus on intra-organizational client and service data and have not evolved to be able to address these more current concerns. Questions that conceptually, at least, seem relatively simple, such as the characteristics or even counts of people who are clients of both mental health and chemical dependency treatment systems, can become quite costly to answer. Developers may have to be satisfied with modest proxies while processes for data system enhancements and integration are still in the future. At the same time, identification of pressing policy and service planning questions can drive future data system development.

Multiple Indicators

It has been emphasized that no indicator can stand alone. In order for an individual PI to have validity and thus utility, it needs to focus on a specific concern. But different aspects of performance in complex systems are interrelated, and measurement of this complexity requires multiple indicators simultaneously tapping key dimensions. The framework offered earlier, along with suggested applications and the scenarios to follow, can serve as a device for planning a comprehensive set.

Participation of a range of stakeholders in the planning process may help to ensure that an appropriate range of indicators is developed, including relevant data sources other than providers and provider organizations. Despite the need to minimize bias and therefore to maximize the objectivity of the data, subjective information (e.g., client satisfaction) is sometimes important to include in a PI system. These data can only be validly and reliably provided from the perspective of clients themselves.

Participation of appropriate stakeholders in development will increase the chances for widespread understanding and consensus around results. Ciarlo et al. (1986) have suggested that measures that allow parallel measurement from several alternative perspectives are to be preferred over singl~perspective measures. Given the political context of a PI system, this argument seems applicable to a PI system as a whole; credible measurement of system performance may need to take into account the view from multiple perspectives.

Concern with inclusion of stakeholders formerly given too little attention (e.g., consumers and families) should not obscure the importance of groups that continue to be central stakeholders for PIs: the direct service providers. They are typically the source of a high percentage of the data used in constructing indicators. Their commitment to data quality may be crucial to its ultimate utility. The inclusion of PIs with direct relevance to the clinical enterprise (e.g., integrated measures of need, treatment process, and service outcome as aspects of responsiveness, efficiency, and effectiveness) will help both to keep data quality high and to maintain the linkage between policy and performance.

Inclusion of multiple indicators does also force consideration of another issue: prioritization of PIs. Since PIs will typically be used, even if indirectly, in a process of decision-making about allocation of resources, there will be explicitly or implicitly some algorithm or other procedure for weighting and integrating findings across multiple indicators. This, too, is an aspect of policy; the more explicit it is, the more impact it can have through a PI system.

Technical Measurement Issues

A second area to which developers of PI systems should be alert concerns issues around the techniques of measurement.

Performance indicators are intended to serve as measures of key aspects of a complex system. All measurements inherently contain some degree of error; performance indicators are no exception to this rule, and they are vulnerable to several types and sources of error. By being attuned to potential pitfalls, by anticipating and avoiding them, practitioners can ultimately use PIs much more effectively and efficiently.

Many of the kinds of issues involved in the use of performance indicators mirror those discussed in great detail in the literature on methodology in research and evaluation. Although particulars of technical discussion typically derive from concerns identified in the development and use of complex, standardized tests, the concepts are applicable even to the apparently simple, unitary ratios used in many PIs. And even if the driving force behind PIs is performance accountability rather than knowledge generation, the essential activity of using information to draw valid inferences is equivalent in the two domains. In addition, there are other issues particular to using data interactively in real-world settings, as use of performance indicators requires. In this section, these technical concerns will be identified and reviewed briefly. References to relevant works on research methodology are provided in the bibliography.

Validity

By their very nature, performance indicators are intended to capture and measure efficiently the most critical dimensions of a system. The first and most important question is therefore whether an indicator validly represents the behavior or characteristic of interest. In designing measures, researchers distinguish several types of validity that are quite relevant to performance indicator system design (Nunnally, 1978; Campbell & Stanley, 1963; Cook & Campbell, 1979).

Predictive validity refers to the degree to which performance on a measure correlates with later actual performance in a domain of interest, i.e., the degree to which a measure predicts real-world performance. This aspect of validity is relevant to indicators designed to measure, for example, capacity, fiscal viability, or need - to those indicators which, rather than showing current performance, are of interest because of their presumed relationship to possible future performance or risk. Choice of indicators with a predictive purpose should be based on their validity for a specific concern as demonstrated in empirically based research literature.

Content validity concerns the degree to which the sources of data for the measure are sufficiently representative of the domain being measured. This type of validity may seem more appropriate to a test of a characteristic such as mathematical ability, in which a specifiable range of information and skills is to be assessed, than to a performance indicator, which is typically a ratio of simple, single variables. However, PI system designers can use the concept of content validity to assess two important questions. First, is a proposed indicator as central as possible to the content area it purports to represent? Since the number of indicators must be limited, each PI should do the best possible job of focusing on the content area of concern. Second, does a subset of indicators, taken as a whole, represent the range of content within the broad range of concern? An important feature of PI development, discussed elsewhere in this report, is that PI systems should if at all possible sample broadly from the universe of possible behavior. Since PIs will typically be used in a context of sanctions and rewards, a broadly representative set of indicators will minimize the risk that organizations may distort performance counterproductively in order to maximize a narrow gain.

Construct validity refers to the issue of whether an indicator is a valid index of the construct it purports to measure. This type of validity goes beyond the notion of content validity and is particularly relevant to abstract concepts for which the content of the domain is not fully specified or operationalized. For example, we are likely to apply such concepts as accessibility or quality to services, and to operationalize them in measures such as percentage of referrals completed or number of critical incidents per given number of patient-days, and so on. An important question may be whether such abstract concepts as these are sufficiently coherent as constructs or whether for measurement purposes they should instead be broken down into smaller, less complex constructs. In other words, are the quantitative data elements employed in the indicator, or set of indicators, broad enough for the concern that generated them, or at least for the range implied by its label?

Establishment of the construct validity of these types of measures would ideally call for first, being able to specify the range of variables related to a given construct; second, determining whether these variables indeed tend to measure the same thing and not several different things; and third, showing empirically that performance on the measure is related to other results that are theoretically predictable, i.e., programs with higher performance on objectively based indicators of quality of services also do generally well on more subjective measures.

Convergent validity represents the extent to which a measure correlates with other measures of the same phenomenon. PI systems put efficiency of measurement at a premium. If a less used but immensely simpler measure is demonstrated to have relatively good convergent validity with a more standard but lengthy measure, it would typically represent a better choice in a PI context. The term *face validity* is used to indicate that a measure appears, on the face of it, to measure what it purports to measure, e.g., because the content meets the criteria of common sense. Many PIs may have good face validity - they may need to have it for practical or political reasons - but consideration of the types of validity discussed above is advised in order to ensure that the effort invested in an indicator does not yield illusory results.

Reliability

Effective application of performance indicators requires that the data be reliable. The general meaning of the word to indicate something that can be depended upon is enhanced in technical usage to signify quantitative reproducibility. PIs are most typically used to compare different entities at a given point in time, or the same entities over time. The user of a measure needs to know that it means the same thing each time it is used, whether by different reporters at the same time or by the same reporter over time.

If an indicator is used to assess performance over time, it is important to be sure that changes in level are indicative of changes in the entity being measured rather than being due merely to error in measurement. A state agency might use in the computation of an indicator the percentage of total un-duplicated caseload in a particular clinical or demographic category. If the same method for including cases and eliminating duplications is not used at two different time periods and used with equivalent accuracy, changes in the indicator will be uninterpretable.

Reliability of equivalent concern when two different sources are expected to provide comparable data. This concern is particularly pertinent when ratings requiring some kind of judgement are used in the production of data for indicators. Staff may rate clients in terms of some level of functioning, or consumers may rate organizations on some measure of satisfaction. In order for these types of data to be used appropriately for any type of comparison, the raters must be using the rating scales in an equivalent way. Clearly, training to ensure consistent use and thus reliable data is necessary in this instance.

Two methods are most commonly used to evaluate the reliability of a measure. A screening instrument, or some kind of self-report measure, e.g., a satisfaction questionnaire, can be administered twice after a sufficiently short interval that true values should not have changed. The equivalence of scores for the two administrations over a number of respondents is a measure of test-retest reliability. Alternatively, a rating measure can be applied by different raters to be same entity; e.g., two case managers equally informed about a client's status can rate strengths or problems. The equivalence of scores across a number of such triads is a measure of inter-rater reliability. For standard measures, these figures are typically reported, although such figures may be valid only for the population and/or settings in which they were derived.

Methods for determining reliability are relatively straightforward and are described in a number of sources (e.g., Rosenthal, 1984). For continuous variables, the intraclass correlation is preferable to a Pearson correlation, since the latter does not take into account equivalence of absolute levels across raters; levels above .70 are typically acceptable. For classification variables, Kappa is most appropriate; values between .40 and .70 may be acceptable for this conservative statistic.

Training is perhaps an obvious necessity for use of rating scales requiring rater judgement, although this training may be incorporated within written instructions. But even when little

judgement seems to be required, reliability can suffer in a number of ways: ordinary human error, possibly magnified by complex computational or logical rules for indicator generation; or well-intentioned but misguided guesswork in an attempt to compensate for missing data. Clear guidelines should be promulgated to specify procedures for generating and reporting data, and face-to-face training will increase the likelihood of data with acceptable reliability. Finally, auditing will provide a mechanism both to encourage careful reporting in the first place, and then to discover the sources of possible losses in reliability.

Reactivity

Social and behavioral scientists endeavor to develop non-reactive measures, i.e., ones that are minimally susceptible to bias or distortion, whether conscious or unconscious.

This goal can be difficult to achieve, particularly with self-report measures, which respondents can often bias in a socially desirable direction. Self-reported alcohol consumption, for example, is significantly lower than objectively measured consumption. In some sense, PI systems are largely organizational self-report data, so users of PI data need to be concerned with this kind of bias. Most data will be objective, rather than subjective, so that opportunity for significant bias would be minimized.

A well-constructed set or indicators, especially in conjunction with credible auditing procedures, would reduce the likelihood that indicators could be falsified. But unclear, loosely specified definitions may allow reporting programs to give themselves the benefit of the doubt. This bias would vary across organizations, and reliability would be reduced. Newman *et al*. suggest that data are more reliable when they are derived from clinical settings in which those who generate the data also use them. Such use reduces bias both through providing another system of verification and through clinicians' greater commitment to accuracy.

Range, Variation, and Sensitivity to Change

For an indicator to be useful, it must be able to detect meaningful variation in the population it measures. If all of the entities (e.g., organization, people) in the population are clustered at one end of the scale of a measure, there may be little chance of observing differences across entities or time. This might be the case, for example, if an outcome measure designed for an average or normal population is used with an extreme population. However, lack of variation across organizations at a single point in time may not necessarily be a problem, if any or all of them may change over time in the context of policy or other external pressures, and this type of change is to be used as an indicator.

Sensitivity and Specificity of Classification Measures

Most PIs are continuous variables; that is, an indicator may have any of a range of values between two numbers, e.g., 0 and 100. Sometimes these may be made up of percentages of items, e.g., cases, meeting certain diagnostic or other classification criteria. when such binary (e.g., yes/no) variables are used, potential problems with misidentification of cases should be considered. The *sensitivity* of a classification measure is its ability to correctly identify cases meeting a criterion (true positive) and is defined as the ratio of true positives to the sum of true positives and false negatives (cases that should have been identified as meeting criteria but were not).

Specificity is the ability of a measure to correctly exclude cases not meeting a criterion (true negatives) and is defined by the ratio of true negatives to the sum of true negatives plus false positives. If standardized classification criteria are used, sensitivity and specificity values may be available in the scientific literature. These values can be calculated for a new measure if a separate, established criterion measure is used in parallel for a sample of cases.

In practice, sensitivity and specificity may be inversely related. A measure, or a particular threshold in a measure, may favor sensitivity at the expense of specificity. That is, it may

emphasize identification of a high percentage of true cases, while at the same time allowing the inclusion of cases that should not be so identified. There is no universally optimal balance. For some purposes, good sensitivity may be important, in order to make sure that as many who should qualify are included; for others, high specificity may be more critical, in order to ensure that only those cases which should qualify are included. Users of PIs based on these types of variables simply need to be aware of the implications of possible false negatives and false positives in interpreting indicator values.

Additional Caveats

Many of the technical and practical issues identified above in relation to planning and development of indicators have direct implications for their use. If a review of potential threats to successful collection of valid and reliable data has been incorporated into planning and discussion, users and audiences will be prepared for realistic interpretation of results. whenever possible, questions about the technical properties of measurements (validity, reliability, etc.) should be addressed on an ongoing basis in order to avoid either over-interpretation or dismissal of findings. Thus, small differences in moderately reliable indicators should not carry significant sanctions, and known limitations of data systems should not be allowed to erode the credibility of significant findings.

In view of all of the potential limitations, a primary caveat is to remember that indicators only indicate. To use a single number as an unquestionable finding with strong sanctions attached may threaten the entire commitment to using PIs. Performance on a particular PI represents a provisional finding, typically requiring further inquiry and verification with different indicators or other data. A system of multiple indicators will allow the effects of random error to be minimized.

There are sources of variation in PIs not due to measurement *per se* that should also be considered in interpreting findings. Many of these are discussed in detail in references on evaluation research methods (Cook & Campbell, 1979; Posovic & Carey, 1989). As with technical issues of measurement, several of these topics are very relevant to using PIs. Even though users do not typically think of themselves as doing evaluation research, they do want to make causal inferences about intervention and change, and they will therefore do well to be aware of possible 'threats to internal validity" - i.e., rival explanations for apparent findings of difference or change.

Maturation refers to "natural" change over time, specifically in people, but the concept *can* be applied to measuring organizations over time. *History* refers to events outside of the formal system that may influence critical behaviors within it. Change may result more from unmeasured environmental influences than from interventions designed to achieve that result. Typically history refers either to trends or to singular events, but cyclical patterns such as seasonal variation may also be a factor. Seasonal variation may not be important when comparing several organizations at any one point in time, but it may be important when comparing performance at any given time to a standard, or to performance at another time. when seasonal variation is known, either through experience, or through empirical reports from elsewhere, raw values can be seasonally corrected.

Selection, differing admissions to programs, *mortality,* or early departure from a program, may bias findings from resulting population differences in groups being compared; certain PIs may need to be population-specific or to be case-mix corrected. *Regression to the mean* can suggest apparently systematic change when only random movement exists; since extreme scores at a given measurement include some component of error, such error may contribute to the extreme score and on average will contribute less in subsequent measurements.

This last issue of extreme scores is taken up with particular fervor in a broader discussion of sources of variation in the TQM literature (Deming 1986). One technique that is particularly helpful in averting over-interpretation in a continuous improvement process is the *control chart.* Most variation in measurement is considered random, i.e., based on influences beyond control and true measurement. Response to variation from a known mean value is therefore indicated only for extreme values. Performance is charted over time, and control limits are derived from the distribution of scores on a particular measure. Performance outside these limits, e.g., the 5% and 95% performance marks, generates a response, while the remaining performance is considered to fall within the normal, non-significant range. The alternative is to chase random variation; since this variation is by definition unrelated to corrective

intervention, efforts to impose beneficial change will be largely futile and therefore ultimately harmful.

This discussion of variation illustrates a significant point in the utilization of PI data. Although most people involved in development and use of PI systems need not have much if any formal training in statistics, this expertise should be available to staff on at least a consulting basis and should influence development of both the indicators themselves and procedures for their interpretation. Issues like statistical power, the capacity to infer true differences from the available data using appropriate tests, are integral to most uses of PIs. A sample may not be large enough to find significant differences. Conversely, failure to apply statistical tests may lead to drawing invalid inferences, for example through not taking into account large standard deviations, the problem the control chart addresses. The substantial investment of resources in PI systems warrants the crucial addition of technical expertise in interpretation. Widespread discussion of these issues in advance, e.g., on when a seemingly large difference is not a difference at all, will help avert the risk of misguided interpretation.

VII. ILLUSTRATIVE SCENARIOS

There are several important themes in this report. First, PIs are based on and used for monitoring and assessment of the implementation of key policies. Second, the relevance and importance of specific policies varies from one situation to another. Third, sound design and development of a meaningful and useful set of PIs requires going through the process of choosing and articulating key policies and identifying questions pertaining to the implementation and effect of these policies.

The Task Force concluded that this document not be a cookbook of indicators, but rather lead the reader through the rationale and steps of the design and implementation of PIs. This chapter underscores these notions through the presentation of illustrative scenarios, the perspective and concerns represented by each and a sample of indicators that can be used to examine relevant policies. There are no "right answers", or "right indicators" for each of the scenarios. The presented sets are examples of potentially meaningful and useful measures.

These scenarios (lettered A-G) were chosen to reflect a wide variety of organizational structures, perspectives, concerns, stakeholders, and policy positions.

A) An SMHA, chosen because state systems are the largest component of public mental health care and perhaps the most important vehicle for change and improvement of service systems.

B) A local area (e.g., county) that contracts with private service agencies for the provision of mental health services. This example is illustrative of the growing trend of public managed care that (unlike the private model) is designed to assure and procure all needed services for its geographically defined population.

C) The perspective and concerns of both consumers and their advocates. This perspective is very important because the effect of services on consumers and their resulting quality of life are what mental health care is all about.

D) A psycho-social rehabilitation service agency. It is an example of an important orientation and a growing segment of the field of mental health care.

E) A private, multiple-location mental health service corporation, which demonstrates the applicability of the model to all sectors of mental health.

F) A private managed care organization, representing another important trend in health care.

G) Assessment of compliance with two legislative requirements of PL99-660.

The PIs are presented purely to illustrate the decision-making process of creating appropriate PIs in a variety of contexts. The Task Force assumed that very few individuals will adopt every indicator of these scenarios. Each scenario is preceded by a brief description of the entity interested in PIs, its perspective, concerns, and the intended use of resulting PIs.

SCENARIO A

Perspective

State mental health agency (SMHA)

Preamble

An SMHA finances and manages the state's public mental health service system through annual service contracts with private, not-for-profit provider agencies. The SMHA conducts its own need assessment, with local input, and allocates funds to service agencies accordingly. Published policies and procedures provide detailed information on management values, policies and expected professional standards. Uniform definitions are established via a detailed data dictionary. Standardization and data reliability are promoted because of the SMHA's emphasis on all three types of comparisons; longitudinal (same agency over time, cross-sectional (agency vis-a-vis similar others) and against *a priori* values (professional standard, own contract commitment, etc.).

Concerns

The SMHA's immediate concerns are, one, getting the most out of every dollar spent, i.e., productivity and cost containment; and, two, having the best client outcomes possible, i.e., increased functional level and improved quality of life. Longer-term concerns involve, one, compliance with legislative mandates; two, meeting community needs; and, three, maximizing consumer satisfaction.

Use

Each agency's service commitments, i.e., who will be served, what services, how much will be provided and with what quality assurance, are spelled out in the annual service contracts. PIs are used to monitor each agency's performance against its own contract and in comparison with similar agencies. Performance that falls either below state average/norms, or below contract commitments triggers technical assistance and, if not improved, risks de-funding of the agency. High performance is publicized and sometimes rewarded by expanded contracts. Thus, the PIs are used to help the state shape its service system in the desired direction, to develop and monitor the service system, as well as reward, or punish, agencies through performance contracts.

Examples of SMHA's Policies and Corresponding Management Questions and PIs

Unit of	Dimension of Performance		

Analysis	Responsiveness	Efficiency	Effectiveness
Consumer Cohort	Policy: Individuals: 1) with severe and persistent mental illness, 2) at risk of developing mental disorders, and 3) at risk of psychiatric hospitalization are the high priority groups for public MH Services. A. Within mental health system Management **question:** What proportion of estimated # of County X's SPMI is being served? **Indicator's name:**	Policy: Commitment to productivity and cost containment, yet financing and providing services that are consistent with need. A. Within mental health system Management **question:** What is the relative cost of serving SPMI? **Indicator's name:** Relative cost of serving SPMI **Indicator's formula:** average expenditure per SPMI average expenditure, all others B. Systems Integration Management **question:** What is the total cost of serving an SPMI? **Indicator's name:** Public cost of serving an average SPMI **Indicator's formula:**	Policy: The bottom line for mental health services is the improvement in consumers' service outcomes; quality of life (QOL), functional level (LOF), standard of living and satisfaction. A. Within mental health system **Managemet question:** Do services increase consumers' LOF? **Indicator's name:** Average change in LOF admission/charge **Indicator's formula:** sum (LOF at discharge-LOF at admission) # discharged consumers B. Systems Integration **Managemet question:** Is linkage to generic services associated with increased QOL? **Indicator's name:** Relative QOL of linked consumers

	Level of County X's treated prevalence of SPMI **Indicator's formula:** registered SPMI in County X estimated #SPMI, County X B. Systems Integration **Management question:** Are generic services available to SPMI? **Indicator's name:** Availability of VR to SPMI **Indicator's formula:**	total costs of MH+generic services to SPMI #SPMI consumers	**Indicator's formula:** average QOL; linked consumers average QOL; clients not linked

	VR slots to SPMI total VR slots		

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
Organization Cohort	Policy: Provide necessary fiscal and human resources to meet the MH needs of high priority consumers and link them to MH and generic services they need. A. Within mental health system **Management question:**	Policy: For greatest effects with limited resources, commitment to providers' and system's productivity and cost containment. A. Within mental health system **Management question:** Does Agency X contain service costs? **Indicator's name:** Agency X's relative cost per unit of service **Indicator's formula:** agency X's cost per unit of service state average cost per unit of service B. Systems Integration	Policy: Preference for contracting with agencies with documented high consumers' service outcomes A. Within mental health system **Managemet question:** How do free standing programs do in terms of consumers' outcomes, e.g. quality of life? **Indicator's name:** QOL of consumers of free standing programs **Indicator's formula:** average QOL; free standing programs average QOL; all other programs B. Systems Integration

	What is the level of mental health financing? **Indicator's name:** County X's level of state, MH funding **Indicator's formula:** County X's state MH $ 10,000 population B. Systems Integration **Management question:** Do homeless SPMI receive MH services? **Indicator's name:** Treated prevalence of	**Management question:** How productive is the inter-system team X/ **Indicator's name:** Average # assessments per team FTE **Indicator's formula:** # children assessed #FTE's in team	**Managemet question:** Are children better off in areas with school-based, mental health programs? **Indicator's name:** Effect of school-based MH programs **Indicator's formula:** # out-of-home placements, per 10,000 children; areas with school-based prog. # out-of-home placements, per 10,000 children; area with no school-based prog.

	homeless SPMI **Indicator's formula:** # homeless SPMI consumers estimated # homeless SPMI		

SCENARIO B

Perspective

A county, or other local area mental health authority responsible for meeting the mental health needs of residents in its geographically defined area.

Preamble

A county mental health authority is charged with assessing its residents needs for mental health and related services, procuring all services and managing the local service system. This authority conducts its own need assessment and contracts with private, not-for-profit service agencies to deliver the services, as specified in annual contracts. Service agencies provide the authority with data on the clients served, services provided, costs and other data agreed upon. The authority also engages in monitoring and evaluation of the total, local service system.

Concerns

The main concerns of the local authority pertain to outreach to persons needing, but not receiving services, assurance of access to services, identification of appropriate costs of service (which could be different for different client groups) and reduced reliance on costly inpatient care.

Use

Findings will be used for "knowledge communication," i.e., feedback to service providers in a way that might improve performance, and "persuasive communication," i.e., using findings

to exert pressure for improved performance. Some findings might also be used for contract negotiations and financial consequences.

Examples of Policies and PIs of a County Mental Health Authority

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
Consumer Cohort	Policy: Provider should meet local commitment to assure service access for PSMD **Management question:** Does the treated prevalence meet the projected county level? **Indicator's name:** **Indicator's formula:** PSMD served in 1991	Policy: Provider's costs will be covered based on consumers served and the respective costs. **Management question:** What is the relative cost of treating PSMD? **Indicator's name:** **Relative cost of a PSMD** **Indicator's formula:** annual avg. Cost per PSMD annual avg. Cost all consumers	Policy: The program should reduce reliance on inpatient care for PSMD **Management question:** Was there a reduced hospitalization rate for PSMD? **Indicator's name:** **Change in hospitalization rate** **Indicator's formula:** Ave inpatient days, per PSMD, 1991 Ave inpatient days, per PSMD, 1989

	estimated #PSMD in county		

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
Organization Cohort	Policy: The PSMD of this county should have access to professional. **Management question:** Are there enough professional to serve the PSMD? **Indicator's name:** **Staff availability for PSMD** **Indicator's formula:**	Policy: Staff effort in serving PSMD should be similar to effort on behalf of other groups. **Management question:** Are staff serving the PSMD as productive as hose serving other groups? **Indicator's name:** Producing of staff service PSMD **Indicator's formula:** Service units to PSMD/FTE service units to other consumers/FTE	Policy: PSMD should decrease reliance on restrictive programs. **Management question:** Do PSMD "graduate" from day programs to clinics? **Indicator's name:** Program "graduates" **Indicator's formula:** #PSMD transfers to clinic, or work in 1991 # PSMD transfers to clinic, or work in 1989

	# professionals/100 PSMD state wide # professionals/100 PSMD		

SCENARIO C

Perspective

Consumers and consumer advocates

Preamble

This group advances the notions that the PSMD should be assured access to services and should have the highest priority for available public resources and publicly financed services. Having worked with legislators on the development of Public Law 99-660, these stakeholders want to see the Implementation of the law and the monitoring of expected results, such as increased functional level of consumers and their $_{empowerment}$ To that end, this group participates, in reviews of SMHAs plans and their accounting for federally financed programs.

Concerns

The consumers are concerned with the implementation of PL99-660 and its intended benefits: meeting mental health and support service needs of PSMD, consumers' empowerment and improvement of their quality of life.

Use

Information derived from indicators will be used to educate all stakeholders, lobby on state and national levels on behalf of the PSMD and deploy political pressure to further advance the agenda of mental health consumers.

Sample Policies and PIs for Consumers

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
Client Cohort	Policy: SMHA must be responsive to the treatment and support needs of its citizens who have severe and persistent mental illness. A. Within mental health Question: Under 99-660, what target has the SMHA set for itself in terms of the proportion of individuals with SPMI who are to be served by community-based services: How is the State doing in achieving this target overtime?	Policy: SMHA resources should be targeted primarily to the populations that are most likely to have to rely on the public system for services, i.e., individuals with SPMI. A. Within mental health Question: Under 99-660, what financial and staffing resources is the State applying that will assist individuals with psychiatric disabilities to gain access to community-based services? Label: Relative per capita expenditures for individuals with SPMI Calculation: Total community-based mental health expenditures Prevalence of SPMI Question: What is the ratio of community-based staff in the 4 core disciplines to individuals with psychiatric disabilities who are registered in these programs? Label: Staff to client ratio Calculation: # individuals w/psych. disabilities enrolled in comm.-based programs # staff of discipline T employed by/ contracting with comm.-based programs	Policy: SMHA sponsored services should produce a positive benefit in the lives of individuals with SPMI and their relatives. A. Within mental health Question: Do clients discharged from mental health programs function at a higher level than those being admitted? Label: Relative discharge functioning Calculation: Average discharge functioning level of clients (-) Average admission functioning level of clients, where (+) score indicates higher discharge functioning. B. System integration Question: As a consequence of 99-660 initiatives, to what extent has the SMHA been successful at activities that empower, train, and involve individuals with SPMI and their families so that their ability to function as citizens is enhanced? Label: Vocation placement index Calculation: Total # SPMI successfully placed in competitive full or part time employment Total # SPMI in vocational training/ preparation programs sponsored by SMHA

	Label: Population coverage. Calculation: Denominator Option 1: 5% of civilian population will be individuals with SPMI. 1.5% X State population = prevalence. Denominator Option 2: Research/other methodology identifies prevalence. <u>#SPMI served</u> (For years X,Y,Z) prevalence For each year the result is compared to the target/goal. <u>B. System integration</u>		

	Question: Are most of the enrolled clients with SPMI simultaneously receiving a support service? Label: Support service access Calculation: #SPMI clients receiving MH+ support services # registered SPMI clients		

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
	Policy: SMHA must be responsive to the treatment and support needs of its	Policy: Clients of the mental health system who have SPMIs should be provided with those services that are most likely to help them function in community settings. A. Within mental health	Policy: The services provided by mental health organizations shall optimize client choice, minimize the use of restrictive care, and show evidence of positive gains for individual clients A. Within mental health

Organization Cohort	citizens who have severe and persistent mental illness. A. Within mental health Question: Do individuals with SPMI represent the majority of the caseload in community-based mental health programs operated/funded by the SMHA? Label: SPMI percent of caseload Calculation: # SPMI registered in program or organization Z Total # of clients registered in organization Z	Question: For every client with SPMI served, what is the ratio of SMHA dollars spent in community-based and inpatient service? Label: SPMI per capita community vs. inpatient dollar ratio Calculation: Total SMHA $ on inpatient services # SPMI served in I/P services Total SMHA $ on comm.-based service # SPMI served in comm.-based service B. System integration Question: Related to 99-660 requirements on access, how quickly is a mental health treatment program able to make a successful linkage to an additional service needed by an individual with SPMI as identified in a treatment plan? Label: Support linkage index Calculation: 1. Average time to linkage Total # days to SPMI clients' contact w/support services R Number of SPMI clients successfully linked to R during Time X	Question: Do clients who are referred to other mental health organizations for treatment exhibit movement from programs that are relatively custodial (e.g. inpatient and residential) to programs that are less custodial (e.g., outpatient and supported housing)? Label: Index of restrictiveness of care Calculation: (# inpatient (# inpatient discharges discharges referred - referred to to supportive intensive residential residential Total number of discharges (-) Value indicates emphasis on referral to settings somewhat more restrictive than other options. The higher the value, the greater the emphasis (+) Value indicates emphasis on referral to settings somewhat more reflective of community integration than other options. The higher the value, the greater the emphasis.

	B. System Integration Question: Under 99-660 how has the State performed in ensuring that individuals with SPMI who gain access to mental health treatment are also assured access to rehabilitation services, case management, housing, and medical care? Label: Linkage to needed services A,B,C...N* * *Service A,B,C...N can be specific treatments, similar program elements, similar organizations,*	2. Complement waiting period Total # days since SPMI clients' Enrollment in MH service A Number of SPMI clients not yet linked to needed support service R		

	clusters of professionals, or individual staff comparisons. Calculations: (each calculated for years X,Y, Service A Number of individuals with SPMI linked to medical/dental treatment Number of individuals with SPMI who are receiving services in community- based setting			

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness

Organizational Cohort (continued)	Service B Number of individuals with SPMI successfully placed in independent or supported housing Number of individuals with SPMI receiving community-based services for whom housing is identified as an issue in the treatment plan		N.B.: This index requires that the other settings to which referrals are made be contrasted in terms of the degree of protection/custody/supervision experienced by the individual with SPMI in those settings. An alternate calculation would use caseload as the denominator and the numerator would contain and contrast numbers of clients enrolled who are receiving various types of treatment and supportive services. For example (# clients (# clients receiving both - receiving outpatient outpatient care care only) and consumer operated self- help experiences) Total number in outpatient caseload B. System Integration As above, but the referral is to a support service, such as vocational, socialization, housing, etc.

SCENARIO D

Perspective

Psychosocial rehabilitation agency

Preamble

A psychosocial rehabilitation agency (PSR) provides comprehensive services to, mostly, PSMD. The services are based on the psycho-social recovery and rehabilitation approach, with major emphasis on the treatment of clients with dignity and on clients' empowerment. The PSR staff assist clients in identifying their own preferences and choices of living arrangement and activities and attempt to meet all service needs. The agency derives its funding mostly from the Department of Mental Health (DMH), but its managers also aim at maximizing revenues from other sources, such as Medicaid, Medicare, Housing and Urban Development (HUD), etc.

Concerns

The major concerns of the PSR agency: 1) meeting clients' service needs, while maximizing revenues, 2) as much as possible, structuring its services according to clients' preferences and choice, 3) whenever possible, engaging clients and former clients as service providers, and 4) delivering services in a manner that maximizes clients' satisfaction.

Use

PIs are used mostly for continuous quality improvement and internal management of the agency. Results are used for providing feedback to direct service providers and for monitoring of the implementation of desired change.

Samples PIs for Psychosocial Rehabilitation Agency

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
Client Cohort	Policy: State DMH will only fund certain types of clients to receive case management. The PSR agency's policy	Policy: PSR agency has a policy of providing residential services to clients per their request for transitional and supported housing. Management question: What is the ratio of supported housing costs to transitional housing costs? Label:	Policy: PSR agency institutes a new vocational program of consumer-provided case management to enhance community tenure. Agency wishes to compare community tenure before and after instituting new program. Management question:

	of serving all clients needing case management means that dollars must be stretched to serve clients not considered fundable by the state. This results in a PSR agency policy aimed at increasing the proportion of State DMH funded clients receiving case management. Management question: What proportion of clients receiving case management are non-funded? Indicator's name: Change in	Ratio of average supported housing costs per client to average transitional housing costs per client. Sample indicator: Ave. Supported housing costs per client avg. Transitional housing costs per client	Is community tenure longer after instituting the new program than before program implementation? Label: Ratio of average length of community tenure after new program implementation to before program implementation Sample indicator: average number of days out of hospital in first year of new program average number of days out of hospital in year before new program

	proportion of clients funded by State DMH to receive case management services. Sample indicator: total # non-funded case management clients total # clients receiving case management services		

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
	Policy: Client satisfaction data indicates client dissatisfaction	Policy: State DMH has a policy of providing equal amounts of funding of staff for residential and vocational services in a PSR agency. PSR agency wishes to know whether the ratio of costs of the two types of services has changed compared to the prior year.	Policy: State Board of Education funds PSR program to offer two types of educational services to members. The first type is for early school levers to help them obtain a GED. The second type is for members who wish to attend community college. PSR program

Organizational	with receiving desired job placements slower than residential, educational, or other program placements. To correct this, the PSR agency institutes a policy of placing clients in desired jobs with speed equal to that occurring for other types of program placements. Management question: Is the average length of wait for vocational placements longer than for all other placements? Label: Change in ratio of staffing	Management Question: Are costs of residential services equal to costs of vocational services? Label: Change in ratio of staffing costs for residential staff to staffing costs for vocational staff. Sample indicator: residential staffing costs vocational staffing costs	policy is to maintain a low dropout rate from both programs. Management question: What percentage of clients drop out from the GED program? Label: Proportion of clients leaving GED program before completion. Sample Indicator: Prematurely % of clients completing GED program

	costs for residential staff to staffing costs for vocational staff. Sample Indicator: residential staffing costs vocational staffing costs		

SCENARIO E

Perspective

A corporation that delivers inpatient mental health services at a number of hospitals

Preamble

A private mental health hospital chain delivers primarily inpatient mental health services at a number of facilities scattered around the country. The corporation must monitor the performance of all hospitals to insure that they maintain a high level of quality of care to preserve accreditation and to protect its professional reputation. The corporation must also insure an adequate level or return on its investment in each hospital. Thus, it must carefully monitor the costs and profits of hospitals and intervene when facilities fail to meet their targets.

Concerns

Assure the delivery of high quality services that improve clients' lives while insuring the financial viability of the organization. Identify under-performing organizations and bring them up to appropriate levels of performance. Identify the best providers and identify what makes them best.

Use

The corporation uses PIs to monitor each hospital's performance against corporate goals and against other corporate hospitals. Performance that is above the corporate goals and better than other hospitals is rewarded with bonuses and is studied for replication by other hospitals. Performance that falls below goals, or the level of other hospitals, triggers interventions to

improve performance. Continued poor performance results in changes in hospital practices, or eventually the closure or sale of the poorly performing hospital.

Examples of a Corporation's Policies, Management Questions and PIs

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
Client Cohort	Policy: The corporation must insure that hospitals are responsive to the service needs of the community to insure that it attracts adequate numbers of clients. A. Within the mental health system Management question: How well are programs meeting the needs of client groups in their	Policy: The corporation must insure that it is spending appropriate resources on all client groups. It must make sure that it is providing adequate amounts of indigent care, but not so much as to jeopardize the financial stability of the organization. A. Within the mental health system Management Questions: What client groups account for the highest and lowest program costs? Indicator's name: Cost per client: Indicator's formula: Expenditures (by client group) Number of Clients (by client group) Indicator's name Staff per client:	Policy: Services are designed to assist clients in terms of clinical outcomes, quality of life measures, client satisfaction, etc. Are the programs meeting these needs? A. Within the mental health system Management question: Which client groups are improving the most with the delivery of services? Indicator's name: Change in client satisfaction by type of service Indicator's formula Change in client satisfaction (by client group) Services delivered Indicator's name: Change in client LOF

	community? Indicator's name: Catchment area's treated prevalence rates for mental illnesses Indicator's formula: Clients Served (by Characteristics) Catchment area population (i.e. Age, Sex, Race, Ethnicity, SPMI, Homeless, etc. B. Systems integration Management Question What percent of clients needing overlay	Indicator's formula: Staff of organizations (by client group) Number of Clients (by client group) B. Systems integration Management Question What is the level of uncompensated care that the organization is providing? Indicator's name: Percent of uncompensated care Indicator's formula: Expenditures for clients without insurance Total expenditures	Indicator's formula: LOF at discharge LOF at admissions (by client group) B. Systems integration Management Question Does provision of services enhance the ability of client to hold supported employment? Indicator's name: Effect of supported employment on LOF Indicator's formula: Improvement in Employment status Services provided	

	services get linked into these services? (I.e., housing, VR, SSI, etc) Indicator's name: Linkage to non-mental health services Indicator's formula: Clients linked into Overlay Services Population needing linked Services		

Samples PIs for Psychosocial Rehabilitation Agency

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
	Policy: Corporate hospitals need to offer an	Policy: To remain financially viable, the hospitals must become productive by controlling staff and capital costs for services	Policy: All hospitals must strive to provide the best quality care possible. Quality care is measured as care that improves the clients' lives. What services and

Organization Cohort	appropriate mix of services that meet the needs of the community. <u>A. Within the mental health system</u> Managemet question: Does the hospital offer an appropriate mix and number of services for the demand in the hospital catchment area? Indicator's name: Hospital's market share of inpatient beds Indicator's formula: <u>Hospital's # of inpatient</u>	<u>A. Within the mental health system</u> Managemet question: Which hospitals have higher or lower costs per client (costs can be determined by staffing rates or fiscal costs) by type of services provided? Indicator's name: Staff to client ratios by type of services Indicator's formula: <u>Staff</u> Services <u>B. Systems integration</u> Management Question: Does working with the local housing agency to locate housing for inpatients decrease the cost per client episode? Indicator's name of Management of housing on cost per client episode Indicator's formula: <u>Cost per client episode for clients w/housing</u> Cost per client episode w/out housing services	hospitals are achieving the best results? <u>A. Within the mental health system</u> Managemet question: Which programs result in the best increase in client level of functioning (LOF) from admission to discharge? Indicator's name: Change in (LOF) from admission to discharge. Indicator's formula <u>Average change in LOF at for service</u> Average change in LOF for all services <u>B. Systems integration</u> Management Question Do clients have better LOF and satisfaction outcomes when they receive housing services form the local housing agency while receiving inpatient care? Indicator's name:

	psychiatric beds # of inpatient psychiatric beds in area Indicator's name: Inpatient bed capacity Indicator's formula: # of inpatient beds offered # of needed inpatient beds in area (as determined by needs assessment) B. Systems integration Management Questions: Are there waiting lists for housing that prevent clients			

	from being discharged?		

SCENARIO F

Perspective

Chief Executive Officer of a not-for-profit, managed care mental health and substance abuse company.

Preamble

This managed care mental health and substance abuse company accepts risk-based captivated contracts from a variety of employers and insurers to manage and provide a full range of psychiatric and substance abuse service to enrollees. Outpatient and day treatment services for mental health and substance abuse clients are run directly by the company. Other mental health and substance abuse services, such as inpatient psychiatric care are contracted to other agencies. Enrollees include a broad mix of employees and their dependents and diverse age groups.

Concerns

The agency's concerns are: 1) the continued viability of the organization, which means staying within the capitated rates for a variety of enrollees, 2) the quality of the services delivered, under the assumption that high quality services reduce long term costs, 3) satisfaction of users (clients and the system) and purchasers of care who are largely employers or insurance companies.

Use

Indicators will be used to monitor how the organization performs against the contract requirements of each entity. They will be used to measure the financial viability of the total organization in an ongoing way. They will also be used to assess: 1) the efficiency of contracted agencies, 2) the effectiveness of services to clients, and 3) clients' satisfaction with services provided.

Sample Policies and PIs for a Private Managed Care Corporation

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness
	Policy:	Policy:	Policy:

Client Cohort	Provide mental health and substance abuse services to the capacity of the contracts to those individuals in the insured population who require services. 1. Managemet question: What is the ratio of the number of persons served to the expected prevalence, by disorder? Indicator's name: Ratio of treated to expected prevalence, by client cohort. Indicator's formula: # of treated	Keep costs of care to clients within capitated rates for each insured group. 1. Management question: What is the relative per capita cost of clients in each insured group? Indicator's name: Expenditure per insured group's client Indicator's formula: monthly, per capita expenditures monthly capitation rate Policy: Maximize outcomes and minimize costs by providing continuity of care. 1. Management question: How prompt is the linkage between discharge from inpatient and enrollment in outpatient services? Indicator's name Continuity of care. Indicator's formula: actual # days between discharge And first outpatient services	Maximize quality service and good business through high client and family satisfaction. 1. Management question: What is the level of client/family satisfaction: Indicator's name: Consumer satisfaction Indicator's formula: consumer satisfaction by client group average consumer satisfaction Policy: Increase quality and reduce financial risks by reducing recidivism/readmission. 1. Management question: What are the readmission rates, by client group? Indicator's formula # readmitted/# discharged children projected # readmissions/# discharges

	children expected prevalence of children 2. Management question: What is the ratio of service use of insured population to use built into contract? Indicator's name: Use rate Indicator's formula: rate of inpatient use expected rate of inpatient use	Contracted # days for linkage	

Unit of Analysis	Dimension of Performance		
	Responsiveness	Efficiency	Effectiveness

Organization Cohort	Policy: Provide, or contract for sufficient staff to insured population. 1. Managemet question: What is the needed level of ambulatory staff? Indicator's name: Needed ambulatory FTEs Indicator's formula: average active cases pre-determined client-staff ratio 2. Management question:	Policy: Services of staff and contracted providers should be as efficient as possible. 1. Management question: Is Provider A's ALOS consistent with others? Indicator's name: Provider A's ALOS Indicator's formula: Provider A's ALOS ALOS all providers 2. Management question: How productive is the staff? Indicator's name Staff productivity Indicator's formula: face-to-face time total time total units of service	Policy: Maximize viability and quality through client satisfaction. 1. Management question: What is the level of client satisfaction by provider organization? Indicator's name: Relative satisfaction of Provider A's clients. Indicator's formula: Provider A's avg. Client satisfaction avg. client satisfaction all providers 2. Management question: Are insurers satisfied with services? Indicator's name: Insurer satisfaction. Indicator's formula: # of appeals this month avg. # of appeals this year 3. Management question:

	What is the number of inpatient beds that the company should contract for? Indicator's name: # of admissions x ALOS 365 x 1.18 (adjustment for .85 Occupancy rate)	#FTEs 3. Management question: How well do we access general-purpose services such as state supported long term care? Indicator's name: Promptness of linkage to long term care.	Are there differential rates of recidivism of contracted providers? Indicator's name: Provider recidivism rates. Indicator's formula: Provider A's recidivism rate Recidivism rate all providers

SCENARIO G

Perspective

State Mental Health Authority (SMHA)

Preamble

SMHAs are bound by both federal and state legal requirements and Public Law 99-660 is one such federal mandate. The 12 requirements of the Law prescribe the development of organized, community-based systems of care for PSMD, articulate the importance of identifying need, assuring adequate fiscal and human resources, assuring access to services, providing case management, reducing use of state hospitals, etc. States can incur fiscal penalties, if they are assessed as non-compliant with the Law.

Concern

Major concerns of the SMHA involve the monitoring of implementation and compliance with PL99-660.

Use

Data derived via PIs will be used for on-going monitoring of the implementation of policies relevant to PL99-660, assessment of compliance and of effects brought about by implementation of the legislative requirements.

Comments

Typical of most legislative mandates, monitoring of policy implementation and compliance with PL99-660 requirements fall in the domain of mostly one dimension of performance, i.e., responsiveness. What is at stake could be described as doing what one is supposed to do, rather than performing efficiently. It should also be kept in mind that translating legislative requirements into PIs is opened to interpretation and can be quite complex. For that reason, and in order not to conflict with any prevailing interpretation, only a brief sample of PIs for assessing compliance with PL99-660 is provided: Requirement #2 for client cohort and Requirement #5 for organization cohort. Both are within the performance dimension of responsiveness.

Sample PIs for Assessing Compliance with PL99-660

Unit of Analysis	Dimension of Performance: Responsiveness
Client cohort	PL99-660 Requirement #2: Specify quantitative targets of SPMI/SED to be served
	Policy: It is incumbent upon each county to ascertain the prevalence of SPMI/SED and enroll and serve as large a proportion of these target groups as possible. Management question: What percent of assessed prevalence of SPMI is registered for service in County X? Indicator's name: Treated prevalence of SPMI; proportion of assessed number of SPMI registered for service in County X Indicator's formula: #registered SPMI in County X Assessed prevalence of SPMI in County X
Organization cohort	PL99-660 Requirement #5: Describe financial resources and staff necessary to implement the requirements of the plan.
	Policy: The public mental health system will assure needed financial and human resources necessary to meet the service needs of SPMI/SED

	Management question: What is the level of state and federal funding of mental health service programs for SPMI, in County X, in comparison with the assessed funding need for those programs in County X? Indicator's name: Ratio of available vs. Needed public funding for mental health programs for SPMI in County X Indicator's formula: current state+federal $ for MH programs for SPMI in County X need for public $ for MH programs for SPMI in County X

Final Comments on administrative Scenarios and Performance Indicators

As with all data-based decision support models, findings from PI Systems should be interpreted with caution. Caution should be exercised with regard to the quality of the data used for the generation of indicators, the reliability of information across organizations or consumer types, the sensitivity and specificity of the measures and the generalizability of PI-based findings. Caution should also be exercised since the interpretation of PIs requires constant review to guard against hastily extrapolating small changes or differences in PIs over time because of the inherent variability of real world performance. This is of particular concern in situations where the numbers of organizations or consumers reflected in the numerators or denominators of these measures are quite small.

VII. CONGRUENCE WITH MHSIP DATA STANDARDS

PIs are a reflection of an organization's values and policies. As a result, the need for indicators and the actual products developed will change over time. An essential feature of the MHSIP has been the development of standards for data content to assure comparability of data across the entire mental health services delivery system and over time. An extensive review process was undertaken by the MHSIP Advisory Group to incorporate data items that were recognized at that time as necessary elements.

The most recent set of standards was published in 1989 in Data Standards for Mental Health Decision Support Systems (NIMH Series FN No.10). These standards have been widely accepted and incorporated into information systems, statistical reporting programs and research.

Data standards were provided in five content areas:

o organization data, which describe the organizations that provide services to the mentally

o consumer/patient data, which describe the characteristics of the mentally ill persons admitted for treatment

o human resources data, which describe the clinical staff of the mental health service provider organization

o event data, derived from activity reports by staff that permit the analysis of the types of service categories received by consumers/patients of mental health organization and the staff who are involved in providing the services

o financial data, which describe the monies received and spent

This task force was charged by the MHSIP Advisory Group to enhance the MHSIP recommended data standards with the design of a system of PIs that can be derived from the content of MHSIP. Many such indicators are possible and are reflected in the various scenarios presented in this report. The MHSIP also recognized that periodic review of data standards is essential to be responsive to changing programs and changing needs.

Meeting the needs of individuals with psychiatric disabilities has changed greatly since the initiation of MHSIP and is likely to continue to change in ways that cannot be anticipated. The MHSIP Advisory Group, in looking to the year 2000 and beyond, has identified a person-based, consumer orientation as offering the most promise for MHSIP. Consistent with this, the process presented in this report can result in a number of useful indicators which draw upon data that are not part of current MHSIP data content. The availability of such data will vary.

The data can be categorized generally as follows:

o data not in MHSIP, but readily available, such as census statistics

o data not in MHSIP, and not readily available to mental health organizations, but which are often routinely generated by other organizations, for example, data relating to consumers and services of other health human Service agencies

o data not in MHSIP and not generally available, such as outcome data

Specific data not currently found in MHSIP content which will be of interest to mental health organizations in generating PIs include:

o Needs data. Population data as well as incidence and prevalence data are examples. Related factors such a suicide rates and child abuse rates are other examples.

o Support services by other health and human service sectors that are actively involved in providing services to persons with mental illness. Primary health care, vocational rehabilitation and housing are examples.

o Outcome measures. These measures include level of functioning, quality of life and consumer satisfaction. Currently, level of functioning is a recommended MHSIP item, but without recommending specific operational definition or instrumentation. Data on consumer satisfaction, waiting period, family involvement and consumer/family resources serve as other examples of useful measures. At this time, however, these items are not usually produced on a routine basis.

The scenarios draw on a number of these examples in order to demonstrate the breadth and depth of current MHSIP elements, as well as promote a person-based orientation.

VIII. REFERENCES

Anderson, H.J. (1991). Sizing Up Systems. *Hospitals,* Oct, 33-34.

Barrett, T.J., Berger, B.L. and Bradley, L.A. (1992). Performance Contracting: The Colorado Model Revisited, *Administration and Policy in Mental Health,* 20(2), 75-86.

Campbell, D.T., & Stanley, J.C. (1963). *Experimental and Quasi-Experimental Designs for Research.* Chicago, IL: Rand McNally and Company.

Ciarlo, J.A., Brown, T.R., Edwards, D.W., Kiresuk, T.J., & Newnan, F.L. (1986). Assessing Mental Health Treatment Outcome Measurement Techniques. National Institute of Mental Health. Series FN No.9. DHHS Pub. No. (ADM)86-1301. Washington, D.C.: US Govt. Printing Office.

Cook, T.D., & Campbell, D.T. (1979). Quasi-experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton-Mifflin.

Fishbein, M., Ajzen, I., & McArdle, J. (1980). Changing the Behavior of ALCOHOLICS: Effects of Persuasive Communication. In I. Ajzen, & M. Fishbein (Eds.), *Understanding Attitudes and Predicting Social Behavior.* (pp.217-242).

Deming, W.E. (1986). Out of the Crisis, Center for Advanced Engineering Study, Cambridge, MIT. Federal Program Performance Standards and Goals Act of 1991 (s.20)

Hadley, T.R., Wilcox, J.T., Rosman, G.R. and Mazar, K. (1983). Performance Standards and Allocation of Funds in Mental Health Programs. *Administration in Mental Health,* 10(3), 155-161.

Jacobs, J.H. and Thompson, J.W. (1986). Lesson from NIMH's Operations Management System for CMHCS. In Windle, Jacobs, and Sherman (Eds.), *Mental Health Program Performance, and Measurement,* DHHS Pub. No. (ADM) 86-1441, 28-36.

Joint Commission on Accreditation of Healthcare Organizations. (1992). Feedback report for obstetrical care and anesthesia indicator data. Oakbrook Terrace, IL: Joint Commission, Department of Outcomes Research and Development.

Kamis-Gould, E. (1987a). Successful Formal Performance Assessment: The Bottom Line is Meaning. In Peterson and Fishman (Eds) - *Assessment for Decisions,* New Brunswick, Rutgers University.

Kamis-Gould, E. (1987b). The New Jersey Performance Management System: A State System and Uses of Simple Measures. *Evaluation and Program Planning,* 10, 249-255.

Kimmel, W. (1983). Performance Measurement and Monitoring in Mental Health; Selected Impressions from Three States. Report to NIMH.

Leff, H.S. and Natkins, J. (1985). Fund Allocation Model and Formulas in State Mental Health Service Systems. Report to NIMH.

Leginski, W.A., et al. (1989). *Data Standards for Mental Health Decision Support,* DHHS Pub.

McLaughlin, J.A., Weber, L.J., Covert, R.W., & Ingle, R.B. (Eds). (1988). New Directions for Program Evaluation: Evaluation Use. San Francisco: Jossey-Bass.

Newman, F.L., Hunter, R.H. & Irving, D. (1987). Simple Measures of Progress and Outcome in the Evaluation of Mental Health Services. *Evaluation and Program Planning,* 10, 209-218.

Nunnally, J.C. (1978). *Psychometric Theory,* Princeton, N.J.. McGraw-Hill. Patton, M.G., 1986. *Utilization-Oriented Evaluation.* Beverly Hills: Sage.

Patton, M.G. (1988). The Evaluator's Responsibility for Utilization, *Evaluation Practice,* 9(2) 5-24.

Peters, T.J. (1987). *Thriving on Chaos,* N.Y., Knopf.

Posara, E.J. and Carey, R.G. (Eds), (1989). *Program Evaluation: Methods and Case Studies,* Englewood Cliffs, N.J., Prentice Hall.

Rosen, A., Miller, V., and Parker, G. (1989). Standards of Care for Area Mental Health Services, *Australian and New Zealand Journal of Psychiatry,* 23:3, 379-395.

Rosenthal, R. (1984). *Essentials of Behavioral Research.* Princeton, N.J. McGraw Hill.

Russell, E.M. and Cole, S.K. (1987). Outcome Indicators: who Benefits. *Scottish Medical Journal,* 32(3), 72-74.

Skinner, P.W., Riley, D., and Thomas, E.M. (1988). Use and Abuse of Performance Indicators. *British Medical Journal,* 297, 1156-1159.

Sorenson, J.E., Zelman, W., Hanbery, G.W. and Kucic, A.R. (1987). Managing Mental Health Organizations with Twenty-Five Key Performance Indicators. *Evaluation and Program Planning,* 10, 239-247.

The State Comprehensive Mental Health Services Plan Act of 1986, (PL99-660).

Tufte, E.R. (1983). *The Visual Display of Quantitative Information,* Chesire, CT. Graphic Press.

Walton, M. (1986). *The Deming Management Method.* New York, Putnam.

Windle, C. (1986). An Orientation to Performance Measurement. In Windel, Jacobs, and Sherman (Eds), *Mental Health Performance Measurement,* DHHS Pub. No. (ADM) 86-1441.

Wholey, J.S. and Hatry, H.P. (1992). The Case for Performance Monitoring, *Public Administration Review,* 52(6), 604410.

1. Currently the Division of State and Community Systems Development, Center for Mental Health Services.