

NIDA RESOURCE CENTER FOR HEALTH SERVICES RESEARCH

Performance Measures in Substance Abuse Treatment

Issue Paper

**Prepared by:
Michael L. Dennis, Ph.D.
Chestnut Health Systems**

January 18, 1999

Author's Note

This paper was written as an initial overview of the topic and the issues involved. It is not a comprehensive review nor does it include every topic or example originally considered in order to maintain a manageable length.

This issue paper was produced for the National Institute on Drug Abuse's (NIDA's) Resource Center for Health Services Research under NIDA contract No. N01DA-5-6050. The opinions expressed herein are solely those of the author and do not represent official positions of NIDA or the U.S. Government. The author wishes to acknowledge comments received on this manuscript from Mark Godley and the NIDA staff, as well as assistance in preparing this manuscript from Joan Unsicker and the Resource Center staff.

EXECUTIVE SUMMARY

In response to near constant pressure from the marketplace to reduce costs, some treatment providers have responded by cutting services and/or quality, and this has raised concerns among sponsors, providers and consumers. Current demands are, therefore, increasingly being linked to a series of “performance standards” and/or are even shifting from “cost” to “cost per some unit of outcome.” Unlike the earlier cost debates between states, managed care firms and providers, there is largely an agreement about the general concept of performance monitoring. The problem is that payers and providers are still negotiating the selection of the measures, where and how much to invest in the system, how to address issues like case mix that can be used to manipulate the results and how to actually use the results to monitor or improve performance. The specific goals of this issue paper are to provide an overview of : a) the forces behind the increasing emphasis on performance monitoring; b) a tentative definition and types of performance monitoring; c) common types of measures; and, d) practical challenges of putting performance monitoring into practice. The paper will draw on both the literature and the author’s extensive experience in developing performance monitoring systems for Chestnut Health Systems and other treatment providers. It will conclude by discussing several recommendations for next steps that could further research and practice in this area.

PERFORMANCE MEASURES IN SUBSTANCE ABUSE TREATMENT

INTRODUCTION AND BACKGROUND

After two decades of managed care, health care reform, and pressure to reduce the cost of substance abuse treatment, attention is increasingly shifting to focus on treatment's effectiveness and efficiency. Cutting costs by cutting services or quality is no longer acceptable. The question is shifting from "cost" to "cost for a given level of performance." While there is largely a consensus on this overall direction, many issues remain unresolved about how to put performance monitoring into practice. The specific goals of this issue paper are to present an introductory overview of:

- The forces behind the increasing emphasis on performance monitoring;
- A tentative definition and types of performance monitoring;
- Common types of measures; and
- Practical challenges of putting performance monitoring into practice.

Much of what is discussed in this paper is a grand work in progress and not in the academic literature. As much as possible the author has, therefore, pointed at some of the key professional forums in the vanguard of the performance monitoring movement.

THE FORCES BEHIND PERFORMANCE MONITORING

The 1990s have seen major changes in the provision of substance abuse treatment: The definitions of substance use disorders have been revised (American Psychiatric Association [APA], 1994); there has been a shift from stand-alone modalities to matching and providing a

continuum of care (e.g., American Society of Addiction Medicine [ASAM], 1994, 1996; Department of Alcoholism and Substance Abuse [DASA], 1995, 1996; Texas Commission on Alcohol and Drug Abuse, 1996); there is continuing pressure to collapse not only alcohol and drug treatment, but substance abuse and mental health into behavioral health treatment systems (Cesare-Murphy, McMahon, & Schyve, 1997; Commission on Accreditation of Rehabilitation Facilities [CARF], 1996, 1998); and public and private funders continue to pressure providers to cut costs and/or provide more services for the same funds (Geraty, 1995; Hall, 1996; Hall & Flynn, 1997). Compounding matters further, virtually every state and federal initiative to reform welfare includes one more performance provision targeted at substance users, particularly mothers or pregnant women. Many include employment provisions that may or may not recognize that chronic substance users often have other disabilities and/or barriers (e.g., AIDS, hepatitis, lack of high school degree, criminal justice record, multiple children, interrupted work histories, no housing) that limit their ability to accomplish this without substantial assistance (e.g., Senay, Dorus, & Joseph, 1981; Dennis, Karuntzos, McDougal, French, & Hubbard, 1993; Dennis, Fairbank, et al., 1995).

Performance monitoring is the common ground that is being encouraged by both accrediting agencies and Congress to resolve conflicts between managed care and providers on lengths of stay and services to be provided, as well as between managed care and clinical researchers who are losing traditional sources of funding through overall cutbacks at the very time the field is demanding more from them (e.g., Agnew, 1996; GAO, 1998a).

DEFINITION AND TYPES OF PERFORMANCE MONITORING

No single definition of performance monitoring has yet reached the level of consensus. For the purpose of this paper, let us define performance monitoring as:

“a system that a) categorizes clients or patients into clinically homogenous subgroups or compares them to a risk-adjusted expectation, b) establishes benchmark expectations of the processes/outcomes and tracks them over time, then c) attempts to optimize these measures by changing the process.”

The categorization or some other form of risk adjustment is a necessary and important step in producing more interpretable and replicable results regardless of whether the analyses are observational or randomized field experiments (Dennis, 1994). Establishing “benchmarks” and implementing continuous measurement procedures are necessary to track changes over time, because not all changes will be planned or controllable (e.g., timing of changes in state regulations or block grants). They are the equivalent of a baseline period in a time series study and have long been the hallmark of performance-based production processes (Drucker, 1995).

The actual approaches to performance monitoring vary considerably. Some repeatedly assess clients to chart progress (e.g., Kazdin, 1993, 1996; Ogles, Lambert, & Masters, 1996), whereas others focus on linking assessment to quality improvement (Salzer, Nixon, & Schut, 1997; Waxman, 1994), managed care applications (Brill, Lish, & Grissom, 1995; Geraty, 1995; Kongstuedt, 1996; Ross, 1997), or the use of report cards and performance indicators] more to evaluate processes at the program level (Dickey, 1995; Kramer, Daniels, & Mahesh, 1996). Also, increasingly more text is targeted at measurement and analytic issues relevant to the use of clinical assessment in practice; this text is aimed at staff in treatment settings rather than at

researchers (Docherty & Streeter, 1995; Hogmann, 1995; Sechrest, McNight, & McNight, 1996; Tonigan, 1995).

Some of the major initiatives already underway in increasing the relevance of these developments to performance monitoring include:

1. Commission on the Accreditation of Rehabilitation Facilities (CARF; www.CARF.org) was among the first to focus on performance monitoring approaches to quality care and released a new manual (CARF, 1998) and other materials to facilitate performance monitoring.
2. General Accounting Office (www.GAO.gov) which has been issuing guidelines (GAO, 1997, 1998a) on the how U.S. agencies are responding to the General Performance and Results Act of 1993.
3. Institute of Medicine (IOM; <http://www2.nas.edu/iom/>) which has just completed a major report (Lamb et al., 1998) calling for changing the process from one in which “research informs policy makers/providers” to a more collaborative approach in which there is a three-way interaction between researchers, policy makers and practitioners - with the practitioners and performance monitoring data driving the researcher’s questions.
4. Joint Commissions on the Accreditation of Health Care Organizations (JACHO; www.JACHO.org) which has mandated that all of its inpatient behavioral health programs start using an accredited outcome monitoring vendor and report risk-adjusted data to the commission so it can be used in tri-annual reviews.

5. National Committee for Quality Assurance (NCQA; www.ncqa.org) which is expanding its Healthplan Employer Data and Information Set (HEDIS) to include some behavioral health requirements and, furthermore, requires that all managed care plans and health networks measure them as part of the new 1999 accreditation process.
6. National Institute on Drug Abuse (NIDA; www.NIDA.NIH.gov) which is attempting to stimulate the serious dissemination and use of new interventions that research has shown to be more effective via its “ Bringing Drug Abuse Treatment from Research to Practice.” (request for applications [RFA] DA-99-007) and National Drug Abuse Treatment Clinical Trials Network (RFA DA-99-004).
7. Substance Abuse and Mental Health Services Administration (www.SAMHSA.gov) which has been changing funding to performance partnerships, setting up state based outcome monitoring systems, studying monitoring in managed care settings and is implementing a common data set across its drug treatment demonstrations as part of its GPRA plan.

It should be noted that many of the above use other related terms like performance standards or outcome monitoring or accountability. What these all have in common is the goal to evaluate and improve a program’s performance over time and/or relative to other programs.

COMMON TYPES OF PERFORMANCE MONITORING MEASURES

While assessment has traditionally been divided by substantive domains (e.g., Dennis, 1998; GAO, 1998b; McLellan et al., 1992), the JCAHO (www.JACHO.org, Cesare-Murphy et al., 1997), others have increasingly divided performance measurement into one of four main types of measures: a) clinical performance, b) health status, c) patient perception of care (i.e., satisfaction), and d) administrative/financial. Table 1 contains examples of some of each type of these measures. Starting in 1999 all JACHO accredited inpatient programs must cover at least 20% of their patients with two to five measures related to the first three types of measures.

Table 1. Illustrative Examples of Measures for Performance Monitoring
<p>Clinical Performance Measures</p> <ul style="list-style-type: none"> % of clients assessed within 48 hours of referral % of eligible and referred clients entering treatment % of clients appropriately matched to level of care (e.g., ASAM 1996 UPPC2) % of clients with psychologically acute problems getting further psychiatric assessment within 14 days of intake % of clients presenting for outpatient treatment sessions % of clients retained in treatment for three months % of clients with successful discharge % of residential discharge clients followed up within 30 days (NCQA requirement) % of residential discharge clients attending three+ outpatient sessions in three months post discharge % readmitted within 12 months of discharge

Table 1 Continued

Health Status/Outcomes

% of days using any drugs
% of days using no alcohol or drugs
% of days in controlled environment
% of days in jail
% of days at work or school
% of days with work or school problems
% of days with family problems
% in at least early remission of substance abuse/dependence (no symptoms in past 30 days) on the GAIN (Dennis, 1998).
Mean GAIN scores on substance problem, substance use, health problem, psychological problem, treatment motivation, relapse and environmental indices.
% with past month ASI (McLellan et al., 1992) composite scores below .25 on psychological, .18 on alcohol and .1 on drug composite (approximately less than weekly problems).

Mean ASI scores on medical, employment, alcohol, drug, legal, family and psychological composites.
Mean RAATE (Mee Lee, 1988) scores on resistance to treatment, resistance to continuing care, biomedical acuity, psychiatric/psychological acuity, social/family environmental indices.
% of women giving birth to normal weight babies without complications requiring post-delivery hospital stays
% of women with 1+ children in protective custody
Outcome status categorized as relapsed in community, incarcerated, inpatient, outpatient, stable in community

Patient Perception of Care

% of clients willing to refer a friend to the program
% who would be willing to come back to the program
Mean score on General Satisfaction Questionnaire (GSQ; Larsen, Atkinson et al, 1979)
Mean subscores (admission process, appointment scheduling, medical staff, and counseling staff) on Client Writes (Barry & Associates, 1998).

Administrative/Financial

Clinical staff to patient ratio
Staff turnover ratio (unique people during year/number of positions)
% of staff meeting accreditation requirements
Average medical coverage (hours per day with on-site medical staff).
Average cost to payer per treatment episode
Average cost to client per treatment episode

PRACTICAL CHALLENGES OF PERFORMANCE MONITORING

General scientific and clinical standards regarding the quality of measurement still apply in performance monitoring, and include issues like face validity, statistical reliability, normed/validating on this or a closely-related population, and a preference for brevity (Dennis, Huebner, & McLellan, 1996). There are also several recurring issues in substance abuse research, including how much to invest in a) measuring substance use (e.g., days, quantity, timeline follow back method), b) measuring substance use disorders (e.g., problems, symptoms, diagnostic measures), c) measuring comorbid problems/treatment that might interact with substance abuse and treatment, and d) validation measures (e.g., urine, collateral reports) or the use of multiple methods. In performance monitoring it is also important to look at several other issues as well:

1. *Defining the population to which a particular measure applies.* Not every measure applies to everyone or every program; it is, therefore, important to define the subset of clients or patients for which a given measure is relevant (the equivalent of inclusion/exclusion criteria in an experimental study). For a given rate, this will be the basis of selecting the denominator and will have one of the greatest impacts on the overall rate and its interpretation.
2. *Defining a successful or satisfactory cut point.* If the above is the denominator, it is often important to define the numerator too. Many research measures are continuous and require the setting of preliminary cut points to interpret and track over time. While the mean can be used, many performance systems track the more directly interpretable percent of people/programs above (or below) an

acceptable cut off (e.g, 85% seen within a week of intake, 80% retained at least three months, 50% in recovery three months post-discharge) - even if they also track the mean.

3. *Risk Adjustments.* Even after selecting a specific population, there is often still variability in severity and/or comorbidity that makes comparisons with other programs questionable without risk adjustment (e.g, two outpatient programs where one has 80% drug court referrals and the other has primarily dual diagnosis clients); this may involve a multivariate predication, comparison group, or a control chart method (Carey & Lloyd, 1995; Wheeler & Chambers, 1992).
4. *Tracking measures over time.* Better performance-monitoring systems will track a person's or program's performance over time, by establishing a benchmark of performance and then seeing if it can be improved upon. When this is combined with norms from other programs it is referred to as a controlled chart (Joint Commission on the Accreditation of Healthcare Organizations [JCAHO], 1997).
5. *Finding measures that are useful for program planning.* Ideally measures selected in performance monitoring should be sensitive to change in the treatment process and/or have fairly direct implications for further change. A general satisfaction measure may be adequate for evaluating a program, but is less useful than a more detailed one (e.g, satisfaction with facilities, staff, intake process) that helps identify more specific areas for program planning.
6. *Using standardized measures meeting above criteria.* There are hundreds of standardized measures used in the field (see Dennis, 1998 for a review), however,

few of them meet the above standards while being sufficiently brief to be integrated into program management. The problem is not with the measures (except perhaps in terms of length), but the lack of norms and analytic work to support them in this application.

Of all these issues, risk adjustment has probably received the least attention, but can totally undermine the validity of any quasi-experimental (and many experimental) comparisons if ignored whether tracking a single program or comparing them. In evaluating the impact of an accelerate methadone admission program, for instance, Dennis, Ingram, Burks, and Rachal (1994) found that the program's six-month retention rate dropped from 91% to 79%. However, most of this change was the result of the grant increasing the percentage of clients on public assistance (from 69% to 88%). Since before and after the grant these public assistance clients averaged 8% to 12% lower retention, the overall rate when down. Once this client variable was controlled for, there was no change in their retention or that of non-public clients (i.e, the change was solely due to a change in the case mix). Unfortunately, there are few formal standards or data on the effectiveness of risk or case mix adjustments. Note that these issues are even germane to experimental studies of effectiveness where the populations are more heterogenous and the odds of the treatment only having an impact in a subgroup increase. Table 2 contains a list of some of common issues that are often considered and currently being explored by several researchers.

Table 2. Common issues considered in case mix adjustments

Age: current age and age of onset

Other Demographics: gender, race, employment status, homeless status

Psychopathology: substance abuse/dependence symptoms and other problems

Motivation: Treatment readiness, resistance/barriers to treatment, pressure to be in treatment, coercion from criminal justice system, stage of change.

Anger/Support: Anger, victimization, other psycho-social stress, social support

Family: Family functioning, cohesion, conflict, substance use

Peers: Peer functioning, substance use/risk behaviors of peers, sexual relationships

Payor: Employer, Veteran’s Administration, public aid, benefit plan limits

Other risk factors: family history of substance use, treatment, special education, social economic status.

LIMITATIONS OF WORK TO DATE

Performance monitoring is spreading quickly and being touted as a potential panacea to the debate on the value of cutting costs by cutting services (vs. waste). Unfortunately, much work to date has some major limitations, including:

1. A focus on “soft” outcomes (e.g., satisfaction, processing time), only a few dimensions of functioning, or psychometrically weak scales (e.g., low or unknown reliability and sensitivity to change) and/or general measures with few direct implications for program planning;
2. Inadequate, unlinked, and unevaluated case mix adjustments (which are essential for non-experimental comparisons) and no, or limited, comparisons groups (either to national norms or the program over time or both);

3. Partial or no linkage with clinical guidelines for diagnosis, placement, treatment planning, or clinical/economic outcomes;
4. Limited range of clients (particularly coverage of adolescents, criminal justice clients, pregnant women, and clients with multiple comorbid conditions) matching to limited programs or services;
5. Limited range of types/levels of treatment, movement along a continuum of care, or high likelihood of re-admissions;
6. Grossly inadequate follow-up rates of public clients (e.g., 20-40%); and
7. No attention to and/or low statistical power (e.g., under 30-50%);

Of the 259 vendors listed by JCAHO (www.jacho.org) at the end of 1998, only 82 list themselves as addressing behavioral health and of these only a handful have any measures targeting substance use. Most of these focus on tobacco or administrative measures (e.g, retention) and none includes a direct measure of dependence/abuse. Of the large Network Performance Measure Templates developed by the end of 1998 (summarized and made available by JACHO as part of the Oryx application materials), three had no substance use measure yet (JCAHO, University of Colorado, University of Wisconsin), only one had goals related to smoking (Foundation for Accountability), and only one (NCQA) had a handful of criteria relevant to substance abuse (e.g., H45 Number of inpatient discharge and average length of stay; H46 Percent of members receiving inpatient, day/night care, and ambulatory services; H47 Number of Readmission for chemical dependency; H48 Outpatient drug treatment utilization). Furthermore, most of the NCQA measures are targeted at performance monitoring of health plans rather than the treatment programs themselves.

RECOMMENDATIONS

Performance monitoring is a promising approach for improving the quality and effectiveness of substance abuse treatment, but an area where much work remains to be done.

Below are several strategic things that NIDA can do to further advance the field.

1. *Help develop norms and benchmarks:* This can be done through the dissemination of norms from earlier national studies that have larger and more generalizable samples (e.g., Gerstein et al., 1997; Hubbard et al., 1989; Sells & Simpson, 1976; Simpson, Joe, & Brown, 1997; Simpson & Savage, 1980) and/or grants to study new systems.
2. *Help develop case mix adjustments:* support grants, workshops, and/or publications directed at reviewing the effectiveness of alternative approaches to case mix adjustments that can be used to improve both statistical power and interpretability in both experimental and non-experimental studies.
3. *Help identify a core set of simple measures:* Consensus panels could be convened to help identify the core measures or at least types of measures that are needed, including simple measures that get at the severity of abuse/dependence (vs. just use) (APA, 1994; Edwards & Gross, 1976; Jellinek, 1960).
4. *Help develop models for evaluating chronic abuse/treatment:* While the majority of people in the treatment system have multiple episodes of care and/or transfer between levels of care - most research still focuses on a single episode and outcomes as though substance use was only an acute condition (Anglin et al.,

1997); performance monitoring of chronic conditions can be very different and has implications for longer-term follow-up and early re-intervention where indicated.

5. *Study the evolution and dissemination of performance monitoring*: performance monitoring itself is an innovation and worthy of study in terms of how it impacts an organization, can be integrated into program planning, and/or spreads to other organizations.

In conclusion, this is a growing area with many challenges. There are gaps in our scientific knowledge, in our application of what we know, and even in our understanding of what providers really need. In short, it is a ripe area for NIDA and others to provide leadership through publications, conferences, and/or funding initiatives.

REFERENCES

- Agnew, B. (1996). Clinical research: Looking for help in unusual places. *Journal of NIH Research*, 8 (7), 21-22.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Society of Addiction Medicine. (1994). *Principles of addiction medicine*. Chevy Chase, MD: Author.
- American Society of Addiction Medicine. (1996). *Patient placement criteria for the treatment of psychoactive substance disorders* (2nd ed.). Chevy Chase, MD: Author.
- Anglin, M. D., Hser, Y. I., Grella, C. E., Longshore, D., & Pendergast, M. L. (1997). Drug treatment careers: Conceptual overview and application to DATOS. *Psychology of Addictive Behaviors*, 11(4), 308-323.
- Barry & Associates (1998). *The Client Writes*. Kokomo, IN: Author.

- Brill, P. L., Lish, J. D., & Grissom, G. R. (1995, September/October). Timing is everything: Prepost versus concurrent measurement. *Behavioral Healthcare Tomorrow*, pp.76-77.
- Carey, R., & Lloyd, R.C (1995). *Measuring quality improvement in healthcare: A guide to statistical process control applications*. New York, NY: Quality Resources.
- Cesare-Murphy, M., McMahonill, C., & Schyve, P. (1997). Joint Commission evaluation of behavioral health care organizations. *Evaluation Review*, 21 (3), 322-329.
- Commission on the Accreditation of Rehabilitation Facilities (CARF; 1998). *1998 Behavioral Health Standards Manual Item (# 5110.22)*, Tucson, AZ: Author(<http://www.carf.org/bhcat97.html>)
- Commission on Accreditation of Rehabilitation Facilities. (1996). *1996 Standards manual and interpretive guidelines for behavioral health*. Tucson, AZ: Author.
- Dennis, M. (1998). *Global Appraisal of Individual Needs (GAIN) Manual*. Bloomington, IL: Chestnut Health Systems (www.chestnut.org/CYT/GAIN).
- Dennis, M. L. (1994). Ethical and practical randomized field experiments. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 155-197). San Francisco: Jossey-Bass.
- Dennis, M. L., Fairbank, J. A., Caddell, J. M., Bonito, A. J., Rourke, K. M., Woods, M. G., Rachal, J. V., Bossert, K., & Burke, M. (1995). *Individualized substance abuse counseling (ISAC) manual*. Bloomington, IL: Lighthouse Publications.
- Dennis, M. L., Huebner, R. B., & McLellan, A. T. (1996). *Methodological issues in treatment services research*. Paper commissioned by the Panel on Effectiveness and Outcomes, Subcommittee on Health Services Research National Advisory Council on Alcohol Abuse and Alcoholism, of the National Institute on Alcohol Abuse and Alcoholism. Rockville, MD: NIAAA.
- Dennis, M. L., Ingram, P. W., Burks, M. E., & Rachal, J. V. (1994). Effectiveness of streamlined admissions to methadone treatment: A simplified time-series analysis. *Journal of Psychoactive Drugs*, 26, 207-216.
- Dennis, M. L., Karuntzos, G. T., McDougal, G., French, M. T., & Hubbard, R. L. (1993). Developing training and employment programs to meet the needs of methadone treatment clients. *Evaluation and Program Planning*, 16 (2), 73-86.
- Department of Alcoholism and Substance Abuse. (1995). *DARTS users manual for treatment counselors, prevention specialists, and computer entry personnel*. Springfield, IL: Author.

- Department of Alcoholism and Substance Abuse. (1996). *Substance abuse treatment regulations*. Chicago: Author.
- Dickey, B. (1995). The development of report cards for mental health care. In L. I. Sederer & B. Dickey (Eds.), *Outcomes assessment in clinical practice* (pp. 156-160). Baltimore: Williams & Wilkins.
- Docherty, J. P., & Streeter, M. J. (1995). Measuring outcomes. In L. I. Sederer & B. Dickey (Eds.), *Outcomes assessment in clinical practice* (pp. 8-18). Baltimore: Williams & Wilkins.
- Drucker, P.F. (1995). Reinventing government. *Atlantic Monthly*, 49-61.
- Edwards, G., & Gross, M. M. (1976). Alcohol dependence: Provisional description syndrome and related disabilities. *British Journal of Addiction*, 85, 357-366.
- General Accounting Office (GAO, 1997). *Agencies' strategic plans under GPRA [Government Performance and Results Act of 1993]: Key questions to facilitate congressional review*, (GAO/GGD-10.1.16). Washington, DC: U.S. General Accounting Office.
- General Accounting Office (GAO, 1998a). *Agencies' Annual Performance Plans Under the Results Act: an Assessment Guide to Facilitate Congressional Decisionmaking*. (GGD/AIMD -10.1.18). Washington, DC: U.S. General Accounting Office.
- General Accounting Office (GAO, 1998b). *Drug Abuse: Research shows treatment is effective, but benefits may be overstated*, (GAO/HEHS-98-72). Washington, DC: U.S. General Accounting Office.
- Geraty, R. D. (1995). The use of outcomes assessment in managed care: Past, present, and future. In L. I. Sederer & B. Dickey (Eds.), *Outcomes assessment in clinical practice* (pp. 129-138). Baltimore: Williams & Wilkins.
- Gerstein, D. R., Ingles, J., Datta, R., Talley, K., Jordan, K., Schildhaus, S., Johnson, R., Rasinski, K., Taylor, J., Bacellar, H., Anderson, D., Phillips, D., Collins, J., Condelli, W., Ciftan, E., & Rohde, F. (1997). *National Treatment Improvement Evaluation Study (NTIES) main findings*. Rockville, MD: SAMHSA Center for Substance Abuse Treatment.
- Hall, L. L. (1996). Report cards accelerate quality and accountability: Impact of managed care on severe mental illness – The role of report cards, consumers, and family members. *Behavioral Healthcare Tomorrow*, 5 (3), 57-61.
- Hall, L. L., & Flynn, L. M. (1997). NAMI's managed care report card. *Evaluation Review*, 21 (3), 352-356.

- Hogmann, A. A. (1995). Measurement sensitivity in clinical mental health services research: Recommendation for the future. In L. I. Sederer & B. Dickey (Eds.), *Outcomes assessment in clinical practice* (pp. 161-168). Baltimore: Williams & Wilkins.
- Hubbard, R. L., Marsden, M. E., Rachal, J. V., Harwood, H. J., Cavanaugh, E. R., & Ginzburg, H. M. (1989). *Drug abuse treatment: A national study of effectiveness*. Chapel Hill, NC: University of North Carolina Press.
- Jellinek, E. M. (1960). *The disease concept of alcoholism*. New Haven, CT: College and University Press.
- Joint Commission on the Accreditation of Healthcare Organizations. (1997). *Oryx: The next evolution in accreditation*. Oakbrook Terrace, IL: Author.
- Kazdin, A. E. (1993). Evaluation in clinical practice: Clinically sensitive and systematic methods of treatment delivery. *Behavior Therapy, 24*, 11-45.
- Kazdin, A. E. (Ed.). (1996). Evaluation in clinical practice (Special Series). *Clinical Psychology: Science and Practice, 3*, 144-181.
- Kongstuedt, P. R. (1996). *The managed health care handbook*. Gaithersburg, MD: Aspen.
- Kramer, T., Daniels, A., & Mahesh, N. (1996). *Performance indicators in behavioral healthcare*. Portola Valley, CA: Institute for Behavioral Healthcare.
- Lamb, S., Greenlick, M.R., & McCarty, D. (Eds) (1998). *Bridging the gap between research and practice. Forging partnerships with community based drug and alcohol treatment*. Washington, DC: Institute of Medicine.
- Larsen, D. L., Atkinson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2*, 197-207.
- McLellan, A. T., Kushner, H., Peters, F., Smith, I., Corse, S. J., & Alterman, A. I. (1992). The Addiction Severity Index ten years later. *Journal of Substance Abuse Treatment, 9*, 199-213.
- Mee-Lee, D. (1988). An instrument for treatment progress and matching: The recovery attitude and treatment evaluator (RAATE). *Journal of Substance Abuse Treatment, 5*, 183-186.
- Ogles, B. M., Lambert, M. J., & Masters, K. S. (1996). *Assessing outcome in clinical practice*. Boston: Allyn & Bacon.

- Ross, E. C. (1997). Managed behavioral health care premises, accountable systems of care, and AMBHA's PERMS. *Evaluation Review, 21* (3), 318-321.
- Salzer, M. S., Nixon, C. T., & Schut, L. J. A. (1997). Validating quality indicators: Quality as relationship between structure, process, and outcome. *Evaluation Review, 21* (3), 292-309.
- Sechrest, L., McNight, P., & McNight, K. (1996). Calibration of measures for psychotherapy outcome studies. *American Psychologist, 51* (10), 1065-1071.
- Sells, S. B., & Simpson, D. D. (Eds.). (1976). *The effectiveness of drug abuse treatment* (Vol. 4). Cambridge, MA: Ballinger.
- Senay, E. C., Dorus, W., & Joseph, M. (1981). Evaluating service needs in drug abuse clients. *International Journal of the Addictions, 16*, 709.
- Simpson, D. D., Joe, G. W., & Brown, B. S. (1997). Treatment retention and follow-up outcomes in the Drug Abuse Treatment Outcome Study (DATOS). *Psychology of Addictive Behaviors, 11*(4), 294-307.
- Simpson, D. D., & Savage, L. J. (1980). Drug abuse treatment readmissions and outcomes: Three year follow-up of DARP patients. *Archives of General Psychiatry, 37*, 896-901.
- Texas Commission on Alcohol and Drug Abuse. (1996). *Agency strategic plan for the 1997-2001 period*. Austin, TX: Author.
- Tonigan, S. (1995). Outcome evaluation: Issues in alcohol treatment outcome assessment. In J. P. Allen & M. Columbus (Eds.), *Assessing alcohol problems. Treatment Handbook Series 4* (pp. 143-154). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.
- Waxman, H. M. (1994). An inexpensive hospital-based program for outcome evaluation. *Hospital and Community Psychiatry, 45*, 160-162.
- Wheeler, D.J., & Chambers, D.S. (1992). *Understanding statistical process control, (2nd ed)*. Knoxville, TN: SPC Press.