

The Role of Monitoring Outcomes in Initiating Implementation of Evidence-Based Treatments at the State Level

Kay Hodges, Ph.D.
Jim Wotring, M.S.W.

This article describes a six-year statewide initiative to help mental health service providers in continuously monitoring the outcomes of youths with serious emotional disturbances who are treated in the public-sector managed behavioral health care system. Participating providers submit outcome data to a state-sponsored evaluator, using the Child and Adolescent Functional Assessment Scale (CAFAS), and receive monthly feedback that identifies youths who are making poor progress in treatment. Additional reports are used to ensure record compliance, monitor at-risk youths, and assist in reviewing the adequacy of treatment plans. In addition, outcome data for closed cases are generated for various types of clients. The consistently poor outcomes for some types of clients have generated a genuine interest among clinical staff in learning and implementing evidence-based treatments. The data for all participating providers were pooled to generate state averages for various indicators so that each provider can compare their site to these benchmarks. State administrators consider the data in generating policy and identifying systemwide needs. The processes that shaped this initiative and that created the providers' investment in continuous quality improvement activities are described. (*Psychiatric Services* 55: 396–400, 2004)

Two of the major challenges facing state administrators are implementing evidence-based treatments and monitoring the quality of care for consumers who are served in the managed behavioral health system in the public sector (1,2). This paper describes a statewide initiative in Michigan, the aim of which was to help service providers develop the capacity to continuously monitor outcomes for the youths and their families whom the providers serve. This initiative, which is now in its sixth year, has created an

environment that has helped poise the state to implement evidence-based treatments in a cooperative and cohesive fashion. It appears to have convinced participating community mental health service providers that evidence-based treatments are important and desirable, lessening the need for the state to initiate change through coercive mechanisms.

Despite the availability of evidence-based interventions, these treatments are not always infused into current clinical practice in the public mental health sector (3). This discrepancy has

resulted in a call to implement evidence-based practices for adults and children with severe mental illness (1,4) as well as to evaluate procedures for transporting these interventions into routine clinical practice after they have been developed in a resource-rich research environment (5).

Tracking outcomes is important, even when evidence-based treatments are used. For many treatments for which manuals are used, only preliminary data are available on mediating variables—which are the specific mechanisms of change—and on moderating variables—which affect generalizability (6,7). Moderating variables are particularly important when transporting evidence-based treatments into applied settings. For example, approximately 70 percent of the families served in Michigan's managed behavioral health system have an income that is below the poverty level, and 75 percent of the households served are single-parent households. Because information on the generalizability of many evidence-based treatments is still limited, it will be important to monitor outcomes of these treatments when they are implemented in routine clinical practice. Furthermore, although monitoring fidelity is critical in ensuring that treatments are faithfully implemented, it is not a substitute for outcome monitoring. A treatment that shows excellent efficacy for one population may not yield a good outcome when applied to a different population or in a different context.

In addition, some of the research on evidence-based treatments has re-


Dr. Hodges is affiliated with the department of psychology at Eastern Michigan University, 2140 Old Earhart Road, Ann Arbor, Michigan 48105 (e-mail, hodges@provide.net). Mr. Wotring is director of programs for children with emotional disturbances at Michigan Department of Community Health in Lansing.

lied on a change in the number of symptoms to demonstrate the treatment's superiority. However, a statistically significant reduction in the number of symptoms does not ensure a reduction in the youth's impairment or in the caregiver's burden (8). Increasingly, there has been a call to require treatment research to include outcome measures that assess the client's everyday functioning in real-world contexts, because these measurements would provide some assurance that the outcomes are clinically significant, as opposed to just statistically significant (9–11). This emphasis on everyday functioning is consistent with the outcome that consumer advocates say that they want for their children. Speaking from a family perspective, Osher (12) stated that families want their children to function better in the natural settings of their communities, specifically, functioning better by living at home, going to school and getting good grades, enjoying friends and activities, and becoming responsible adults who are able to live independently. This mandate translates to determining whether treatments result in better day-to-day functioning in real-world contexts.


Outcomes monitoring for youths in Michigan

For the past six years, clinical supervisors of children's services at 27 community mental health service providers throughout Michigan have actively participated in a voluntary project—known as the Level of Functioning Project—that monitors treatment outcomes for youths aged 6 to 18. Although most participants in the Level of Functioning Project are clinical supervisors, information services experts and quality assurance staff members also participate in the project. The goal of the project is to improve the quality of care by helping each community mental health service provider collect clinically meaningful outcome data, which can then be used to promote continuous quality monitoring. The Michigan Department of Community Health contracts with an independent university-based evaluator to assess outcome for youths whose care is provided by

community mental health service providers but is funded by monies distributed by the state. This initiative began when the director of children's services for one of the community mental health service providers sought out the evaluator's assistance in assessing outcomes for the youths served by the agency using the Child and Adolescent Functional Assessment Scale (CAFAS) (13). This collaboration between the evaluator and one community mental health provider resulted in a pilot project that was well received and eventually published (14). State administrators became aware of the collaboration and saw the potential to empower all



*Helping
providers use
outcome data to improve
their programs and services
was viewed as a win for
the state and for the
consumers.*



the community mental health service providers to use outcome data to improve their programs and services.

Helping providers use outcome data in this way was viewed as a win for all parties: for the state, because its moral and regulatory obligation to the state's consumer constituency would be advanced (2); for the consumers, because of the potential for more effective services (12); and for the providers, because they could master a set of skills that would put them in more control of their own fate as managed care infused the public sector. However, the state recognized that candid self-evaluation would be difficult to achieve if the

providers feared repercussions. Thus a "hold harmless" policy was adopted at the inception of the project to allow for a grace period that would permit agencies to identify and address weaknesses. Because the clinicians were so willing to improve their skills, there has been no reason to end the grace period. Provider-specific data were given only to the provider and the provider did not suffer any negative consequences from the data.

The state also realized that high-quality care could not be accomplished by fiat but rather by a professional culture in which commitment, program initiative, and skill building are highly valued. It followed that the project should encourage these qualities by giving the participants an active role in the shaping the process, with the participants determining what would be helpful and, in essence, collectively guiding the project. Only in this way would the participants assume the responsibility for generating meaningful questions about services and for solving any identified problems. In effect, all parties offered guidance to this dynamic process, but the clinical supervisors were at the center, with the state administrators and the evaluator acting as facilitators and lending their expertise. The following sections describe the project as well as the changes that took place over time in generating and using the data.

Practical operations of the project

The project is inclusive; any community mental health service provider can participate. It was agreed that all participants would be invited to quarterly Level of Functioning meetings, in which data requested by the participants would be presented.

To collect data for the project, all youths served by the participating community mental health service providers were evaluated at intake, every three months after intake, and at the time of exit from services. The CAFAS (13) was chosen as the primary outcome measure, partly because of its psychometric characteristics (15–19) and partly because of the positive experience that one of the community mental health service

providers had with the scale (14).

The CAFAS essentially consists of behavioral descriptions, for example, expelled from school, that are arranged into four levels of impairment—severe, moderate, mild, and no or minimal—across eight domains of functioning, which form subscales—that is, school or work, home, community, behavior toward others, moods and emotions, self-harmful behavior, substance use, and thinking. The rater, who is typically the treating clinician, reads the items in each subscale, beginning with the severe impairment items, until a description of the youth's functioning is found. The youth's score on each subscale is determined by the level of impairment under which the item appears: severe, 30; moderate, 20; mild, 10; and no or minimal, 0. The CAFAS identifies specific behaviors that need to be addressed and generates a score for each subscale and a total score, which is the sum of all eight subscales. Each subscale also has an accompanying list of strengths and goals.

Each community mental health service provider collects data electronically and then selectively exports data to electronic files that are given to the evaluator each month. The software that is used for data input and export was developed by the author of the CAFAS. The software collects CAFAS data as well as other information that can be used to interpret the outcome data, such as child and family descriptors, treatment and services delivered, and other outcome information. For each CAFAS input, the user can generate an assessment report, which includes a graphic display of change in CAFAS subscale scores over time, and a treatment plan, which lists the target problems that have been identified by the CAFAS as well as the strengths and goals, which are listed separately for each CAFAS subscale. SPSS data and syntax files are generated by the program's export utility, which greatly expedites the evaluator's ability to analyze the data and provide feedback in a timely manner. The data can also be exported in other forms, such as Access, Excel, and ASCII, thus permitting each com-

munity mental health service provider to generate reports and conduct analyses.

Early in the project, attention was devoted to ensuring the integrity of the data collected. Three issues were identified: collecting data at the level of the individual item endorsement on the CAFAS, establishing and maintaining reliability, and using the CAFAS in a way that made it clinically useful. The scores on the CAFAS subscales are based on item endorsements, not on a clinician who assigns scores. Global scores generally do not perform well as measures of outcome, likely because of their vulnerability to respondent bias (20). Thus data were collected at the level of the individual item endorsements, which determines the CAFAS subscale scores, rather than at the level of the subscale scores or total score. These item endorsements are kept in the youth's record to ensure that the scores are based on accurate information and that the information is available for caregiver review.

The second issue was to try to guarantee interrater reliability. Although the CAFAS consists of behavioral items, there can be a lack of reliability in the use of such seemingly simple descriptive words as "aggressive." Explicit rules for scoring the CAFAS are contained in the *CAFAS Self-Training Manual* (21). A training-of-trainers model was implemented for the project, with each provider using the same criteria for satisfactory reliability (22) and keeping records to verify the reliability of the raters at their site. Booster trainings were done to ensure continuous reliability.

Unless outcome data have real use, the thoughtfulness that goes into scoring the measure is generally very limited, thus jeopardizing the reliability of the data. Usefulness is easy to establish with the CAFAS, because clinicians typically view it as a treatment-planning tool (15,23). Change in the youth's functioning over time can easily be tracked with the CAFAS. These procedures for ensuring the data's integrity have remained stable for the past six years; however, the types of data analyzed and how the data are applied have evolved over time.

Evolution of the data analyzed

At the beginning of the project, the data presented at the Level of Functioning meetings were from the aggregated state database. However, individual participants soon requested separate results for their program, which they could then compare with the state averages. At that stage, the data were aggregated reports that described characteristics and diagnoses of the youths served, the services and collaborations provided, and the outcome for youths whose treatment had terminated. Cluster analysis was conducted to identify the types of clients that most often used the providers' services, which permitted each community mental health service provider to learn about its effectiveness with different types of clients and to compare its own statistics with the pooled state database. These data sparked interest in what was actually happening with individual clients, because in some cases the results for individual sites did not correspond to the site's treatment philosophy—for example, was it true that many severely impaired, conduct-disordered youths were receiving only individual outpatient therapy? When such questions arose, community mental health service providers requested data at the level of the individual client. These data told the provider whether each client met the criteria for each of five outcome indicators developed for the CAFAS (22).

Although this level of data analysis was informative, it was limited because it only described youths who had exited services. The participants now wanted to receive data on clients who were still in treatment. The providers were hoping that using the data would lead to improved outcomes for clients who were still receiving services. This desire led to the current practice of generating monthly reports for each community mental health service provider that gives outcome-to-date data for each client. The following questions are answered: Is every youth receiving a quarterly outcome evaluation in a timely fashion? Which clients are currently making poor progress? Which clients are at high risk of poor response to treatment, out-of-home

placement, high service use, or high service cost? Monthly, the sites also receive eight Excel charts, which present aggregated data describing their caseloads.

In addition to these monthly reports, the project continues to produce aggregated reports for each site and for the state. The Level of Functioning provider reports are PowerPoint presentations that contain 45 Excel charts describing the characteristics and outcomes for all clients admitted during the fiscal year as well as providing separate statistics for all cases closed during the fiscal year. A comparable PowerPoint depicting the pooled state data results is also provided for each site.

Evolution in the use of data

The community mental health service providers use the data for a variety of purposes. The aggregated data are used to generate presentations for the community mental health service provider's oversight boards, to provide information on outcome and continuous quality improvement activities to accrediting bodies, to report mandated CAFAS data to the state, and to evaluate programs within the agency. One site issues press releases after quarterly Level of Functioning meetings to inform the community about the site's areas of excellence. Some providers have become data champions, a term that describes persons who become advocates for using data for rational decision making (24).

Some participants became champion managers, in that they showed other providers how to use the data to improve the clinical supervision of individual cases and the oversight of delivery of clinical services (25). When data at the level of the individual client became available, the ability to use data to influence clinical flow and decision making increased. Specifically, this information helped the community mental health service providers accomplish several goals: prevent unnecessary restrictive placements; maximize benefits to clients; provide supervision and support to staff, especially on behalf of youths who are not progressing well or who are at risk; manage agency resources

responsibly; and hold staff accountable for records compliance.

State administrators have also actively used the data. When the state had to draft regulations about client eligibility for enhanced intensive services, the state requested input from the Level of Functioning project. At the Level of Functioning meeting, participants generated numerous proposals for determining eligibility for enhanced services. These proposals were translated into algorithms by the evaluator, who then produced data for each of the proposed algorithms for each provider. Because the analysis was able to determine that there was little difference in how the various proposals af-

■
*The
response
to the state's
training-related activities
has been overwhelmingly
positive.*
■

fected providers, conflict between the providers was avoided.

The results for the state database have been analyzed to determine outcomes for various types of clients and patterns of comorbidity for various types of clients (26,27) and to identify predictors of poor outcome (28). These data revealed that large numbers of youths had conditions for which evidence-based treatments exist, such as mood disturbance and behavioral problems at home and at school. In response, the state has taken numerous action steps, including convening a working committee of various stakeholders, such as family and consumer advocates, to guide steps for disseminating evidence-based treatments; implementing a training program in cognitive-behav-

ioral treatment of depression, which includes the unique feature of offering six months of weekly supervisory consultation; planning similar training on parent management training; and grant-seeking activities to obtain monies to study how best to disseminate evidence-based treatments in the public mental health setting.

Thus far the response to the state's training-related activities has been overwhelmingly positive. The providers were very interested in sending representatives to the cognitive-behavioral training, and the participating clinicians gave all aspects of the training program very positive reviews, especially the ongoing six months of weekly supervisory consultation. Because all participants have baseline data, providers are able to determine whether the training is associated with improved outcomes for depressed youths.

The state has also used the data to identify local community-based programs that have exceptional outcomes for highly impaired youths with serious emotional disturbance. Propensity analysis was used to determine whether one such program was superior to a comparison group derived from the Level of Functioning database. Propensity analysis permitted matching the two groups on nine pretreatment variables, including severity of impairment at intake (unpublished manuscript, Hodges K, Grunwald H, 2003). Local programs, such as the one identified by this propensity analysis, can provide consultation to other providers who are hoping to improve services for youths for whom there are no tailor-made evidence-based treatments available. This endeavor would certainly be enhanced if an empirical study of the identified program could be undertaken for the purpose of delineating the mediating variables responsible for the therapeutic change.

Various stakeholders are also discussing other strategies for improving care, including how to promote the retention and recruitment of well-trained practitioners and how changes in financing mechanisms could enhance services for families. A commitment to using outcome monitoring for continuous quality im-

provement remains, in large part, as a result of the active interest of clinicians. Owning and understanding data appear to have promoted a sense of control and mastery that helped to mitigate feelings of apprehension about being evaluated, which often accompany reporting of outcome data for accountability purposes. It is hoped that providing an opportunity for clinicians to actively participate in outcome monitoring and to generate solutions to the problems discovered will result in improved services to families. Continued outcome monitoring will help us know whether that is indeed the case. ♦

Acknowledgment

The Michigan Department of Community Health provided funding for this project.

References

1. Carpinello SE, Rosenberg L, Stone J, et al: New York state's campaign to implement evidence-based practices for people with serious mental disorders. *Psychiatric Services* 53:153–155, 2002
2. Mowbray CT, Grazier KL, Holter M: Managed behavioral health care in the public sector: will it become the third shame of the states? *Psychiatric Services* 53:157–170, 2002
3. Goldman HH, Ganju V, Drake RE, et al: Policy implications for implementing evidence-based practices. *Psychiatric Services* 52:1591–1597, 2001
4. Torrey WC, Drake RE, Dixon L, et al: Implementing evidence-based practices for persons with severe mental illnesses. *Psychiatric Services* 52:45–50, 2001
5. Schoenwald SK, Hoagwood K: Effectiveness, transportability, and dissemination of interventions: what matters when? *Psychiatric Services* 52:1190–1197, 2001
6. Kazdin AE: Bridging the enormous gaps of theory with therapy research and practice. *Journal of Clinical Child Psychology* 30: 59–66, 2001
7. Weersing VR, Weisz JR: Mechanisms of action in youth psychotherapy. *Journal of Child Psychology and Psychiatry* 43:3–29, 2002
8. Angold A, Costello EJ, Burns BJ, et al: Effectiveness of nonresidential specialty mental health services for children and adolescents in the "real world." *Journal of the American Academy of Child and Adolescent Psychiatry* 39:154–160, 2000
9. Kazdin AE, Weisz JR: Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology* 66:19–36, 1998
10. Weisz JR, Hawley KM, Pilkonis PA, et al: Stressing the (other) three Rs in the search for empirically supported treatments: review procedures, research quality, relevance to practice and the public interest. *Clinical Psychology: Science and Practice* 7:243–258, 2000
11. Kazdin AE, Kendall PC: Current progress and future plans for developing effective treatments: comments and perspectives. *Journal of Clinical Child Psychology* 27: 217–226, 1998
12. Osher TW: Outcomes and accountability from a family perspective. *Journal of Behavioral Health Services and Research* 25: 230–232, 1998
13. Hodges K: Child and Adolescent Functional Assessment Scale, 2nd Revision. Ypsilanti, Mich, Eastern Michigan University, 2000
14. Hodges K, Wong MM, Latessa M: Use of the Child and Adolescent Functional Assessment Scale (CAFAS) as an outcome measure in clinical settings. *Journal of Behavioral Health Services and Research* 25:325–336, 1998
15. Hodges K: Child and Adolescent Functional Assessment Scale, in the Use of Psychological Testing for Treatment Planning and Outcomes Assessment, 3rd ed. Edited by Maruish ME, Mahwah NJ. Lawrence Erlbaum Associates, in press
16. Hodges K, Wong MM: Psychometric characteristics of a multidimensional measure to assess impairment: the Child and Adolescent Functional Assessment Scale. *Journal of Child and Family Studies* 5:445–467, 1996
17. Hodges K, Wong MM: Use of the Child and Adolescent Functional Assessment Scale to predict service utilization and cost. *Journal of Mental Health Administration* 24:278–290, 1997
18. Hodges K, Doucette-Gates A, Kim CS: Predicting service utilization with the Child and Adolescent Functional Assessment Scale in a sample of youths with serious emotional disturbance served by center for mental health services-funded demonstrations. *Journal of Behavioral Health Services and Research* 27:47–59, 2000
19. Hodges K, Kim CS: Psychometric study of the Child and Adolescent Functional Assessment Scale: prediction of contact with the law and poor school attendance. *Journal of Abnormal Child Psychology* 28:287–297, 2000
20. Hodges K, Gust J: Measures of impairment for children and adolescents. *Journal of Mental Health Administration* 22:403–413, 1995
21. Hodges K: The Child and Adolescent Functional Assessment Scale Self-Training Manual, 2nd Revision. Ypsilanti, Mich, Eastern Michigan University, 2000
22. Hodges K: CAFAS Manual for Training Coordinators, Clinical Administrators, and Data Managers, 2nd Revision. Ypsilanti, Mich, Eastern Michigan University, 2002
23. Koch JR, Lewis A, McCall D: A multistakeholder-driven model for developing an outcome management system. *Journal of Behavioral Health Services and Research* 25: 151–162, 1998
24. Hodges S, Woodbridge M, Huang LN: Creating useful information in data-rich environments, in *Developing Outcome Strategies in Children's Mental Health*. Edited by Hernandez M, Hodges S. Baltimore, Md, Paul H. Brookes Publishing Co., 2001
25. Barchholtz PR: The use of outcome data on the individual client level to manage treatment, in *The 13th Annual Research Conference Proceedings, a System of Care for Children's Mental Health: Expanding the Research Base*. Edited by Newman C, Liberton CJ, Kutash K, et al. Tampa, Fla, University of South Florida, Louis de la Parte Florida Mental Health Institute, Research and Training Center for Children's Mental Health, 2001
26. Hodges K, Xue Y, Wotring J: Use of the CAFAS to evaluate outcome for youths with SED served by public mental health. *Journal of Child and Family Studies*, in press
27. Hodges K, Xue Y, Wotring J: CAFAS outcomes for children with problematic behavior in school and at home served by public mental health. *Journal of Emotional and Behavioral Disorders*, in press

Change of E-Mail Addresses for Authors and Reviewers

Authors of papers submitted to *Psychiatric Services* and peer reviewers for the journal are reminded to visit Manuscript Central at <http://appi.manuscriptcentral.com> and keep the contact information in their user account up to date. Because the system relies on e-mail communication, it is especially important to keep e-mail addresses current. If you have questions about the information in your user account, contact the editorial office at pscentral@psych.org.