ORIGINAL PAPER

Factors Influencing Consumer Outcome Monitoring in Implementation of Evidence-Based Practices: Results from the National EBP Implementation Project

Doug Marty · Charlie Rapp · Greg McHugo · Rob Whitley

Abstract *Objective* This study reports participating agencies use of a single fidelity dimension, Outcome Monitoring, during the 2-year National EBP Implementation Project. *Method* The study involved an emerging theme qualitative approach for analyzing the implementation of fidelity achieved for five EBPs in 49 sites across eight states. *Results* Twenty-seven percent of the sites reached a high level of implementation fidelity. *Conclusion* Results suggest that four interrelated factors contributed to whether agencies successfully implemented Outcome Monitoring-agency data collection methods, agency culture, practitioner skill, and how well "consumer outcomes" are defined.

Keywords Mental illness · Evidence Based Practice (EBPs) · Community Mental Health Center · Outcome monitoring · Case management

The last 20 years has witnessed a broad-based call for outcome monitoring as a centerpiece of organizational functioning. At the policy and system level, the Government Performance and Results Act of 1993 (Office of Management and Budget 2007) requires that all federal agencies measure the results of their programs and restructure their management practice to improve these results. In a parallel fashion, there is a significant movement in human service management toward client outcome-based management methods (Poertner and Rapp 2007; Rapp and Poertner 1992). Studies have shown that an outcome orientation of managers leads to increased service effectiveness in mental health (Harkness 1997; Gowdy and Rapp 1989). This has led Patti (1985) to argue that effectiveness, interpreted as client outcomes, should be the "philosophical linchpin" of human services organizations.

While monitoring of EBP implementation through fidelity measurement is important (Bond et al. 2000), outcome monitoring is also a critical part of organizational feedback and a powerful managerial tool for improving performance. Implementation monitoring focuses on the structure and process of service, where outcome monitoring informs on the results of those efforts. A review of 20 years of organizational research found that 58% of the 64 studies showed statistically significant positive effects of feedback, 41% showed mixed results, and 1% showed no effect (Alvero et al. 2001). The act of collecting information not only provides data that can be used as a corrective mechanism to improve programs. The process itself generates human energy around the phenomenon being measured. Feedback of that information motivates behavior toward that performance (Nadler 1977; Taylor et al. 1984; Taylor 1987).

The foundation of evidence-based practices is client outcomes. The decision to implement an evidence-based practice is based on its ability to help clients achieve the highest rates of positive outcomes. Therefore, one key component of the implementation of an evidence-based practice is the careful monitoring and use of client outcome data. Consumer outcomes are those aspects of consumers' lives that we seek to improve through the delivery of the evidence base practice. Some outcomes are the proximal

D. Marty (✉) · C. Rapp
School of Social Welfare, Office of Adult Mental Health,
The University of Kansas, Twente Hall, Lawrence,
KS 66044, USA
e-mail: dmarty@ku.edu

G. McHugo · R. Whitley
Department of Psychiatry at the Dartmouth Psychiatric
Research Center, Lebanon, NH, USA

result of an intervention, such as getting a job through participation in a supported employment program, whereas others are more distal, such as improvements in quality of life due to having a job. Furthermore, some outcomes are concrete and observable, such as the number of days worked in a month, whereas others are more subjective and private, such as satisfaction with vocational services.

Effective agencies need to pay attention to both the implementation process (e.g., through fidelity ratings) and *how* consumers respond (e.g., consumer outcomes) (Poertner and Rapp 2007). One without the other only illuminates part of the picture. Monitoring these outcomes may also be a morale boost for practitioners, inasmuch as it recognizes their hard work done on behalf of consumers. This is an additional benefit to identifying service elements that may need attention. A simple but accurate phrase to remember: what gets measured, gets attention, gets done.

For these reasons, "Outcome Monitoring" was one dimension included in the General Organization Index (GOI) that was used as part of the National Evidence-Based Practice Implementation Project (SAMHSA 2007). The GOI consisted of a total of 12 items designed to measure the organizational practices believed to be facilitative of EBP implementation. In addition to the GOI, outcome monitoring was one of twenty-six dimensions used to evaluate implementation of the EBPs. The Project studied the effects of implementing five EBPs across the nation: Assertive Community Treatment (ACT), Family Psychoeducation (FPE), Illness Management and Recovery (IMR), Integrated Dual Diagnosis Treatment (IDDT), and Supported Employment (SE).

This paper reports on the experience of 49 sites in developing, implementing and using outcome monitoring systems as part of their EBP implementation efforts. The paper explores the barriers and inhibitors that were confronted, and the strategies and facilitative conditions that were contributory.

Method

Sample

This paper is based upon data gathered as part of the National Evidence-Based Practices Project (see Torrey et al. 2003; McHugo et al. 2007, for further details). The project involved structured qualitative and quantitative observation of implementation of five psychosocial EBPs in 52 sites across eight states. Period of observation lasted 2 years with the first year being known as the implementation phase and the second as the sustaining phase. Sites were assisted in implementation by an intervention package that included the monthly provision of training by a

consultant trainer (CAT) and further provision of written materials. The intervention was delivered most intensely during the implementation phase and tapered off during sustaining. Overall oversight and co-ordination was provided by the Psychiatric Research Center-Dartmouth Medical School (PRC). In addition to PRC support, CATs and local data gatherers (known as Implementation Monitors: IMs) were directly associated with university based research and training centers (RTCs) in four states. In these states, the RTC provided support and training to CATs and IMs.

Data Collection

Implementation data were collected over 2 years by IMs during monthly site visits. This involved the observation of training sessions, leadership meetings, team meetings and informal conversations with staff, families and consumers. Workers were also shadowed and 6-monthly interviews were conducted with the EBP program leader and the CAT. Core questions posed at these interviews include (i) perceived reasons for success in areas performing well (ii) perceived reasons for shortcomings in areas doing not so well (iii) perceived influence of stakeholders (senior staff, consumers, trainers) on implementation. The IM took detailed notes during/after each observation and formal interviews were taped and transcribed. None of the data sources were weighted over the others. The aim of collecting data from such diverse sources and using diverse methods (i.e., formal interviews and participant observation) was to ensure that multiple perspectives were obtained. This is known as 'triangulation' in qualitative research giving the analyst a chance to cross-check views and opinions. This diminishes bias associated with reliance upon one source or method of data (Malterud 2001). We approached the data in such an open-manner. One aim of the project was to discern the level of fidelity achieved at implementation sites to an EBP model that an expert panel had previously considered efficacious. This panel consisted of national leaders in mental health services research, who judged efficacy through systematic literature review of randomized controlled trials and quasi-experimental studies. Models with consistent and weighty evidence were deemed evidence-based practices (Mueser and Drake 2005).

Formal fidelity reviews were conducted every 6 months by the IM throughout the 2 years of the project. GOI ratings were assigned values on a likert scale from (1) to (5). A (1) indicating agencies that were not incorporating the item and (5) denoting those who reached maximum level of fidelity. In this paper, we do not focus on overall fidelity scores or every measure, but scores of the fidelity item specific to

outcome monitoring. Each fidelity scale had a question relating to outcome monitoring. Within this item, a (1) was defined as "No outcome monitoring occurs". The highest value (5) was defined as "Standardized outcome monitoring occurs quarterly and results are shared with EBP practitioners". The implementation monitor and trainer made separate, independent fidelity ratings. A very high inter-rater reliability was indicated with scores ranging from ICC.99 ($N = 52$) for ACT to ICC.89 ($N = 48$) for IDDT.

Coding National Data

The collection and analysis of the qualitative data was guided by the recommendations of Miles and Huberman (2002) regarding complex multi-site analysis. They suggest that development of inductive categories as they arise from the data is suitable for small single-site case studies, however is fraught with perils and pitfalls for large multi-site studies, where there are differences in context, processes, and data-gatherers. They thus suggest the development of a priori categories that are manageable and transferable across sites. These generic guidelines were followed in formulating design and analysis in the present paper, with all codes being developed a priori.

Firstly, a deductive coding schema was pre-designed by the national project coordinating center in order to capture influential processes and dynamics relating to implementation. The coding schema consisted of 26 possible dimensions (see Table 1), outcome monitoring being one. After importing all transcribed interviews and field-notes into ATLAS.ti qualitative software, implementation monitors were instructed to code any data that related to one of the 26 dimensions. In order to enhance coding rigor, IMs participated in monthly conference calls throughout the project and attended annual meetings to learn, discuss and clarify any process or technical issues arising. To further enhance reliability, the PRC requested that IMs submit examples of coded data monthly for internal review and feedback. When coding was complete, IMs were instructed to write a standardized final site report of approximately 50 pages that summarized implementation processes and dynamics in a qualitative and quantitative fashion. Forty-nine out of the 52 original EBP sites provided a final 2-year implementation report.

One table contained in the final report was a Dimensional Summary of the Implementation Process (see single site example in Table 2). The display used separate rows for each of the 26 predetermined coding dimensions to identify major themes within dimensions that occurred throughout the project as facilitators, strategies, or barriers to implementation. A theme was considered to be a thread of activity or condition that was "salient, prominent,

conspicuous, or non-ignorable". Facilitators identified evidence of factors that helped EBP implementation but were not intentionally developed as a result of implementation. Strategies identified evidence of intentional actions seeking to help EBP implementation. Barriers identified evidence of factors that impeded EBP implementation. IMs were asked to write about these processes giving appropriate examples and commentary. Key stakeholders involved with the theme were also identified. This display was chosen as the major information source for analysis of EBP outcome monitoring implementation process in this paper. The display contained the most relevant and appropriate data vis-à-vis outcome monitoring.

Data Analysis

The research team assembled all 49 completed site reports in Fall 2006. The dimension display 'outcome monitoring' was subsequently isolated at each site and end-point GOI fidelity scores noted. Sites were ranked from highest to lowest 'outcome-monitoring' end-point fidelity. This grouping allowed all four investigators to identify common themes associated with agencies achieving higher fidelity to outcome monitoring and those with lower fidelity.

Once this ranking was complete, each investigator independently assessed the dimension displays to identify themes that appeared to enhance or enable outcome monitoring as well as themes that inhibited implementation. Regular conference calls were held between research team members to discuss independent assessment and develop consensus. By the end of analysis, we reached a high degree of congruence with regards to thematic identification. Examples of prominent themes found in the display included the presence or absence of support, proficiency, knowledge, and technology. The research team agreed that identified themes could be subsumed into two broad superordinate categories that significantly affected outcome monitoring at the sites: these were: (i) actions taken within the agency, (ii) external resources available to the agency. Once these had been established a 2×2 table emerged as a coding heuristic (Table 3). Upon completion of the table, data were reviewed once more. Each identified theme was labeled and placed within the appropriate cell in the table by the primary investigator. The research team reviewed the results and agreed to the efficacy and accuracy of the data reduction.

Findings

This section is organized by grouping the sites into one of three groups and reporting the aggregate findings for high

Table 1 Operational definitions of the 26 implementation dimensions

Prioritization

Attitude	Evidence of expressions of support/opposition for/against the implementation/offering of the EBP
Understanding	Evidence of whether understanding of the EBP is/is not present/being sought
Mandate	Evidence that the MHA does/does not require that this EBP be implemented/offered
Money	Evidence of the presence/absence of financial backing for the implementation/offering of this EBP

Leadership

Responsibility	Evidence that a program leader has/does not have responsibility/authority to implement/offer this EBP
Leadership skills	Evidence that leadership skills for EBP implementation/delivery are/are not present or being sought
Plan enactment	Evidence that there is/is not planning for the implementation of the EBP
Engagement	Evidence that there are/are not efforts to build active support among other stakeholders for implementing/offering this EBP
Plan sustaining	Evidence that there is/is not planning for the sustained offering of this EBP
Change culture	Evidence that there is/is not a culture that is open to the kind of changes needed to implement/offer this EBP

Work force

Staffing	Evidence that there is/is not attention focused on meeting the staffing requirements of the EBP
Personnel action	Evidence that personnel problems that detract from EBP implementation/delivery are/are not addressed
Mastery	Evidence that the skills needed by practitioners to offer the EBP are/are not present/being sought
Training	Evidence that training in the EBP is/is not being offered
Supervision	Evidence that clinical supervision of the EBP is/is not being offered

Work flow

Staff meetings	Evidence that a meeting structure that supports the EBP is/is not present
Documentation	Evidence that documentation structures that support the EBP are/are not present
Support staff	Evidence that support staff functions that support the EBP are/are not present
Physical environment	Evidence that the physical environment needed to support the EBP is/is not present
Policies	Evidence that policies that support the EBP are/are not present

Reinforcement

Penetration	Evidence that measures of program penetration are/are not collected/used to improve the EBP
Outcome monitoring	Evidence that outcome monitoring methods are/are not collected/used to improve the EBP
Fidelity	Evidence that measures of fidelity are/are not collected/used to improve the EBP
Reward	Evidence that success is/is not celebrated/acknowledged to reinforce the EBP
Credential	Evidence that programs are/are not credentialed to reinforce the EBP
Feedback	Evidence that feedback from practitioners/consumers is/is not solicited/used to monitor/improve the EBP

Table 2 A single site example for the dimension "outcome monitoring"

Facilitator	Strategy	Barrier	Net theme
CEO and other senior staff valued outcomes highly. The agency had measured outcomes in the past and received awards for their work.	Admin eventually installed and used outcome-tracking software provided by XXX (COMP). They shared outcomes with CMs on IDDT team.	Senior staff and PL had difficulty deciding on outcomes to track and how to track them. When they received COMP from the XXX, they initially had trouble installing and using it.	Outcomes to track decided on by late imp. COMP installed and outcomes reported to IDDT CMs by mid sust.

achievers, moderate achievers, and low achievers. The groupings reflect the degree to which sites met fidelity standards for consumer outcome monitoring. High achievers embraced and used outcome monitoring in their practice much more readily than low achievers. When implementation themes were extracted from the data and coded, high achievers identified many more themes undertaken to enhance or enable outcome monitoring than did low achievers. Table 4 displays the percentage and number of identified themes used to enhance outcome monitoring versus the percentage of those themes identified as inhibitors to implementation.

Table 3 Final coding heuristic

	Agency actions	External resource actions
Enhancers	Pre-existing agency infrastructure conducive to collecting outcome data	Third-party intervention mandates (e.g. state-level requirements, etc.)
	Agency leadership supports EBP	Additional technical support available to and used by agencies (e.g.: RTCs)
	Local support for aggregating and using data	Database(s) made available for agency use
	Agency champion advocates/implements outcome monitoring	ACT and SE EBPs better defined with easier outcomes to collect/use
	Agency developed/acquired and used new outcome monitoring tool	
Inhibitors	Agency staff have no desire to collect	State systems don't aggregate data in meaningful ways
	Staff feel tracking outcomes is too hard or not worth the trouble	Outcomes poorly defined by some EBPs
	Agency does not use data to manage EBP	State systems do not collect relevant outcome data
	Agency leadership does not support data collection	State systems do not produce timely reports
	Lack of resources (time, money, computers, software, secretaries etc)	

Table 4 Percentage of themes identified by achievement level and type

	High achievers	Moderate achievers	Low achievers
Enhancing themes (%/# of themes)	97% (28)	64% (25)	36% (23)
Inhibiting themes (%/# of themes)	3% (1)	36% (14)	64% (40)
Total	100% (29)	100% (39)	100% (63)

High Achievers

Of the forty-nine agencies reporting, 13 (27%) had a fidelity score of "5" for outcome monitoring at the end of the project. Ten of these agencies (77%) were located in two states. What differentiated these sites from others may be attributed, in part, to the support agencies received from networks outside of the participating agency. Both of these states provided extensive support and technical assistance concerning outcome monitoring to the EBP sites in conjunction with a RTC. This included software programs for use in recording consumer outcomes and producing reports, as well as easily accessible technical assistance that provided guidance and clarity to the agency when relevant questions emerged. The technical support included CATS who were actively involved in assisting programs with defining, collecting and using consumer outcomes on site as well as RTC support staff that provided technical assistance. Both states provided supervisors with a 2-day training focused on consumer outcome-oriented management (Carlson and Rapp 2007).

The nature of the desired consumer outcomes for the specific EBP seems to be influential. Nine of the twelve sites achieving a rating of 5 were implementing either ACT or SE. ACT and SE had the highest average outcome monitoring item scores (3.92; 3.56 respectively) across all EBP's at the end of the 2-year period. Common to both is that there is a single dominant measure of success (prevention of psychiatric hospitalization for ACT and competitive employment for supported employment) that has long standing eminence in mental health, enjoys broad based support from a wide variety of constituencies, and has been routinely measured for many years. In contrast, IDDT, FPE, and IMR involve slightly more nebulous and challenging proximal outcomes that are not binary in nature. These include progress in stages of change (IDDT), emotional support and family problem-solving skills (FPE), and knowledge of mental illness and using medications as prescribed (IMR). These outcomes are not routinely collected, lack consensually accepted instruments, and most involve consumers self-report which add a dimension of complexity to data collection. In fact, one state experienced much debate over the 'proper' outcomes for IMR.

Another key characteristic for agencies meeting high fidelity was the presence of a local champion. There was a person who took enthusiastic responsibility for implementing outcome monitoring at these sites. In some cases, this may have been an agency administrator but in the majority of cases it was the program (team) leader (PL). Program leaders at high fidelity sites ensured the collection, dissemination, and use of consumer outcomes for the team. This appeared to be a function of a positive attitude, motivation, skills and support. Often program leaders were given significant support and tools by outside resources (such as instruction or data base structures) but that was not always the case. In the three cases meeting high fidelity

where external supports were not prevalent, outcome monitoring was achieved through internal agency resourcefulness and fortitude. In these situations, the program leader saw value in collecting outcomes and took the initiative to develop their own system and method. This is evidenced by display data for one of these agencies which had little or no success during the implementation phase of the project but "During the early sustaining phase, the PL used the toolkit materials, taught herself ACCESS and developed an outcome database" on her own. This led to significant improvements in the sustaining phase.

It appears that familiarity with collecting and using data as a management tool also played a role. Over half of the sites reaching maximum fidelity had some experience with using data as a tool of management in their everyday work capacity prior to the project. Although data may not have been specific to consumer outcomes or used in a timely manner to enhance service delivery, it provided practitioners with familiarity to the process. At one high fidelity site, an IM suggested that "due to the site's financial structure, staff were accustomed to collecting and reporting outcomes". Familiarity appeared to have made it more comfortable for some agencies to use consumer outcome data for program development.

Moderate Achievers

There were 14 agencies (37%) which met a moderate level of fidelity (a rating of 3 or 4). Half of these sites (7) were in two states that relied heavily on a centralized statewide information management system. Although these systems mandated the collection and storage of consumer data, the usefulness of the effort relative to EBP appears questionable. Three major issues arose that inhibited effective consumer outcome monitoring in this regard. First, the centralized systems in both states were designed to collect data every 6 months: "… the system is programmed to collect and compare data in 6 month intervals, which is a maximum score of 4, which is a barrier for higher fidelity". These systems also had a slow "turn around" time for generating reports and disseminating these to agencies. Therefore, even if these agencies received reports, there were outdated and lacked usefulness for program development. Another IMP for a different site wrote: "[The] outcome system served as a reporting requirement, not an outcomes system, because of the slowness in developing and demonstrating the relevance of outcome reports". Lastly, centralized systems did not meaningfully aggregate data to be useful at the local level. For example, some systems produced reports using agency specific data, but no centralized system produced reports containing data for a specific team or practitioner. At a minimum, it is at the

team level where consumer outcome monitoring begins to become useful for clinical enhancement (Poertner and Rapp 2007). A report from Virginia also found a lack of clinical utility from a statewide MIS which hindered implementation (Blank et al. 2004). These tendencies of centralized systems greatly interfered with effective implementation. This was true even when PLs demonstrated skills and enthusiasm for collecting and using consumer outcome data.

Even with adequate support and resources, some agencies did not meet a high fidelity standard. One-third of agencies scoring in the moderate level of fidelity were located in the two states that provided both local and RTC technical assistance as described earlier. These agencies chose not to take advantage of the support. In most cases, the PL just simply refused to collect or distribute consumer outcomes throughout most of the project. It appears that when this occurred, PLs felt too busy to collect the data or simply did not see the benefit in doing so. An IMP from one site writes, "Despite repeated requests and offers of assistance, the PL failed to collect client outcomes". At this site, it was not until late in the project that the administration finally stepped in and forced the PL to collect data, but even then the data was not being shared with practitioners in a timely manner.

For agencies achieving moderate success, the presence of inhibitors affected their scoring. Unlike high fidelity agencies, programs in this range recorded approximately half as many inhibitors as positively coded factors used for effective implementation. The vast majority of these inhibitors were related to the lack of desire to implement. More specifically, there appeared to be either internal resistance to implementing outcome monitoring within the agency or key staff (at the local or state level) showed a lack of effort or integrated strategic planning to support implementation. The apparent lack of a synergy between support services and the implementing agencies became a contributing factor to their modest level of fidelity.

Low Achievers

There were 22 participating sites that were assessed as meeting either no or minimal standard for outcome monitoring (scores of 1 or 2). For the vast majority of theses sites (18) limited technical assistance with outcome monitoring appeared to be provided by the CAT. This meant that in numerous instances, the quality of implementation rested mainly on the shoulders of the intervention team. As one IMP noted regarding a low-fidelity site, "[PL] asked CAT for TA [technical assistance] on developing outcome measures, TA not received". This is contrary to those agencies where high fidelity was witnessed. At the majority of high

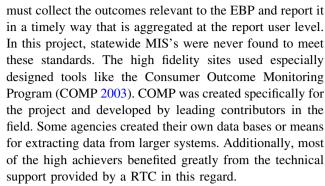achiever sites, an RTC provided guidance as well as project oversight which helped move the project forward.

Another overwhelming characteristic for agencies demonstrating low fidelity was the fact that they tended to be either FPE or IMR sites. All of the FPE (4) and 75% IMR (9) sites fell into this category. This meant that over half of the sites with low fidelity come from these two practices. For these sites, the data suggest that agencies just did not track specific consumer outcomes. A possible reason for this may be that participants in these EBPs reflected a small subgroup of the total population served and agencies were unwilling to incorporate new data collection standards into their practice. It is also evident that agencies implementing these practices often were confused as to what outcomes to collect or agency staff were simply not interested in collecting data.

For the majority of FPE and IMR sites, the data suggests that "no new outcomes specific to [these practices] were collected or analyzed". In part, this may have been due to the lack of clarity around what participating IMR sites were expected to collect. At one site, the IM states that "the state EBP champion initially led an effort…to get [input on] outcome data from several [agency] sources" but the process never yielded effective outcome development. The data suggests that the group decided to wait for the development of a statewide system in the hopes that that effort would add clarity and be useful for the practice. The state effort did not materialize over the course of the project. For IMR sites where some measure of outcome monitoring fidelity was met, the PL at the specific mental health center initiated the discovery and use of an instrument to measure recovery. This suggests that guidance from the CATs as well as the IMR and FPE toolkits lacked the necessary specificity to direct agencies toward what consumer outcomes should be collected or how to use them.

As evidenced in sites that demonstrated a moderate level of fidelity (but much more pronounced for sites with low fidelity), the emergence of inhibitors to implementation greatly effected adherence to outcome monitoring. For sites meeting low standards for fidelity, there were nearly twice as many coded inhibitors as enhancers contained in the display. It appears that the lack of positive actions taken necessary for implementing effective consumer outcome processes combined with the large number of unresolved problems experienced at these sites plagued them throughout the project.

Discussion

The data suggests four interrelated factors affecting consumer outcome monitoring. The first concerns the Management Information System (MIS) itself. In short, it must collect the outcomes relevant to the EBP and report it in a timely way that is aggregated at the report user level. In this project, statewide MIS's were never found to meet these standards. The high fidelity sites used especially designed tools like the Consumer Outcome Monitoring Program (COMP 2003). COMP was created specifically for the project and developed by leading contributors in the field. Some agencies created their own data bases or means for extracting data from larger systems. Additionally, most of the high achievers benefited greatly from the technical support provided by a RTC in this regard.

Second, the organizational culture in which the EBP operates seems influential. Organizations with a history of outcome data collection found EBP outcome monitoring a natural extension. Organizations that did not value such data tended to lack the adequate development of resources to collect and maintain data or train practitioners and managers on how to use data for enhancing performance. The data clearly reflects that a greater number of inhibitors surfaced in agencies where outcome monitoring was not valued. This phenomenon significantly effected implementation.

Third, the skills, attitudes and appreciations of the manager powerfully affected implementation. The high achievers had program leaders who appreciated the power of systematic feedback concerning consumer outcomes. They took the lead and devoted considerable energy to developing the necessary MIS and, more importantly, routinely fed the data to their teams, and helped them bring meaning to the numbers in a way that program improvement could emerge. Two of the states that contained 10 of the 13 high achievers provided a two day supervisory training to program leaders on outcome based management which included content on the power of information, data, and feedback; practice in reading and interpreting outcome data reports; and skills in converting data to program improvement efforts (Poertner and Rapp 2007). Such a training program may be a necessary complement to the presence of an adequate MIS.

Lastly, consumer outcomes need to be clearly defined and disseminated for each EBP. It is no surprise that ACT and SE sites met fidelity standards much more often than other EBPs. These EBPs have very specific outcomes which have been clearly defined and integrated into the practice. There is an urgent need to clarify appropriate outcomes for the other EBPs'. Without such clarity, standardization of the outcome monitoring process is impossible. Further, once specific outcomes are clearly defined, instruction on how to use the data in practice should be spelled out and readily available to all who are interested in the practice.

The experience of the 49 sites would suggest that statewide MIS do not function well as outcome monitoring tools. The data contained in these systems lacked

timeliness and relevancy or the report sent back to the agency is too protracted to be useful. Also, reports are not aggregated and formatted sufficiently to facilitate use at the team level. Yet these statewide systems offer the promise of more complete data sets enabling more sophisticated analyses. One solution to bridge the gap between promise and performance would be for the state to dedicate a team (e.g. program staff and programmers) to work exclusively on developing and producing meaningful reports for the field. This group would need to be shielded from other demands for data from state officials, legislators, etc. This function could also be contracted to an RTC.

The second recommendation relates to the need for resource development at the local level. Agencies need to provide and encourage training of supervisors in the value and use of outcome data to enhance performance (Carlson and Rapp 2007) as well as insure sufficient technical and data entry support availability for maintaining the outcome monitoring program. When these were in place, the EBP sites were able to in fact monitor outcomes.

Limitations to the Study

Continued research would be useful to further inform on the use of consumer outcomes in practice. The data used for this study was collected as part of a large national project. A contributing factor to the limitations of the study was the very nature of that project. The project was charged with exploring 26 separate dimensions of EBP implementation of which, outcome monitoring was only one. Findings were not extrapolated from a study which specifically focused on consumer outcomes. The study is a broad-brush approach where the investigators achieved breadth more than depth. Future case studies may be able to flush out more finely nuanced details. Lastly, even with the consistent oversight of the PRC, there is the possibility of observer bias that cannot be dismissed. It would be surprising, given the numerous data sources as well as the broad nature of the inquiry, that the findings would be significantly influenced however.

References

Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985–1998). *Journal of Organizational Behavior Management, 21*(1), 3–29.

Blank, M. B., Koch, J. R., Burkett, B. J., & Goldman, H. H., (2004). Less is more: Virginia's performance outcomes measurement system. *Psychiatric Services, 55*(6), 643–645.

Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H.-W. (2000). Measurement of fidelity in psychiatric rehabilitation. *Mental Health Services Research, 2*(2), 75–87.

Carlson, L. S., & Rapp, C. A. (2007). *Basic supervisor's training manual*. Lawrence: The University of Kansas School of Social Welfare.

Consumer Outcome Monitoring Package (COMP) (2003). [On-line] Available: http://www.research.socwel.ku.edu/ebp/. Accessed May 2007.

Gowdy, E., & Rapp, C. A. (1989). Managerial behavior: The common denominators of effective community-based programs. *Psychosocial Rehabilitation Journal, 13*(2), 31–51.

Harkness, D. (1997). Testing interactional social work theory: A panel analysis of supervised practice and outcomes. *Clinical Supervisor, 15*(1), 33–50.

Malterud, K. (2001) Qualitative research: Standards, challenges, and guidelines. *Lancet, 358*, 483–488.

McHugo G. J., Drake R. E., Whitley, R. et al. (2007) Fidelity outcomes in the national implementing evidence-based practices project. *Psychiatric Services*.

Miles, M. B., & Huberman, A. M. (Eds.) (2002). *The qualitative researcher's companion*. Thousand Oaks, CA: Sage Publications.

Mueser, K. T., & Drake, R. E. (2005). How does a practice become evidence-based? In R. E. Drake, M. R. Merrens & D. W. Lynde (Eds.), *Evidence-based mental health practice: A textbook* (pp. 217–241). New York: W.W. Norton.

Nadler, D. A. (1977). *Feedback and organization development. Using data based methods*. Reading: Addison-Wesley.

Office of Management and Budget (2007). [On-line]. Available: http://www.whitehouse.gov/omb/mgmt-gpra/gplaw2m.html. Accessed May 2007.

Patti, R. J., (1985). In search of purpose for social welfare administration. *Administration in Social Work, 9*(3), 1–14.

Poertner, J., & Rapp, C. A. (2007). *Textbook of social administration: The consumer-centered approach*, 2nd edn. New York: Haworth Press.

Rapp, C. A., & Poertner, J. (1992). *Social administration: A client-centered approach*. White Plains: Longman Publishing Group.

SAMHSA (2007). Using general organizational index for evidence-based practices. [On-line]. Available: http://www.mentalhealth.samhsa.gov/cmhs/communitysupport/toolkits/community/goi/default.asp. Accessed May 2007.

Taylor, M. S. (1987). The effects of feedback on the behavior of organizational personnel. *Administration in Social Work, 11*(3/4), 191–203.

Taylor, M. S., Fisher, C. D., & Ilgen, D. R. (1984). Individuals' reaction to performance feedback in organizations: Control theory perspective. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resource management* (Vol. 2, pp. 81–124). Greenwich: JAI.

Torrey, W. C., Finnerty, M., Evans, A., & Wyzik, P. E. (2003). Strategies for leading the implementation of evidence-based practices. *Psychiatric Clinics of North America, 26*, 883–897.