



Published in final edited form as:

Psychiatr Serv. 2015 July 1; 66(7): 674–676. doi:10.1176/appi.ps.201400531.

Practical Monitoring of Treatment Fidelity: Examples From a Team-based Intervention for People With Early Psychosis

Susan M. Essock, Ph.D.,

New York State Psychiatric Institute and the Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York City.

Ilana R. Nossel, M.D.,

New York State Psychiatric Institute and the Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York City.

Karen McNamara, L.C.S.W.-C., Ph.D.,

Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland.

Melanie E. Bennett, Ph.D.,

Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland.

Robert W. Buchanan, M.D.,

Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland.

Julie A. Kreyenbuhl, Pharm.D., Ph.D.,

Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland.

Sapna J. Mendon, L.M.S.W.,

New York State Psychiatric Institute, New York City.

Howard H. Goldman, M.D., Ph.D., and

Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland.

Lisa B. Dixon, M.D., M.P.H.

New York State Psychiatric Institute and the Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York City.

Abstract

Mental health programs can address many components of fidelity with routinely available data. Information from client interviews can be used to corroborate these administrative data. In an application of this approach, data from these sources indicated that a team-based intervention for

Corresponding author: Susan M. Essock, Ph.D., se2176@columbia.edu, essocks@nyspi.columbia.edu.

Conflicts of interest: Drs. Bennett, Dixon, Essock, Goldman, McNamara and Ms. Mendon may be part of training and consultation efforts to help others provide the type of services for individuals with first episode psychosis provided as part of the RAISE Connection Program described in this report. These individuals do not expect to receive any personal compensation for any such training efforts, rather such efforts would be carried out as part of the work done for their employers and, in that sense, compensated. In the past two years Dr. Buchanan has served on the following: Advisory Boards: Amgen, EnVivo, Roche; Consultant: Abbott, BMS, EnVivo, and Omeros; DSMB: Pfizer. The remaining authors reported having no conflicts.

people experiencing early psychosis was implemented as intended, including program elements related to shared decision-making as well as a range of evidence-based clinical interventions.

Keywords

Fidelity; first episode psychosis; schizophrenia; performance measurement; state mental health systems

Fidelity measures serve multiple stakeholder groups: payers, trainers, supervisors, clients, and families. Payers want to know if they are getting what they are paying for. Trainers and supervisors want to know if training succeeded and whether clinical staff members are implementing interventions as intended. Clients and families want to know if services are effective and can be expected to promote outcomes they care about (e.g., school/work/friends/health). Fidelity measures are critical to understanding how good outcomes are achieved, replicating successful programs, enhancing efficacy, and measuring performance over time.

This report presents a practical approach to measuring fidelity used in the Recovery After an Initial Schizophrenia Episode (RAISE) Connection Program, a team-based intervention designed to implement evidence-based practices for people experiencing early psychosis suggestive of schizophrenia (1, 2). The project was carried out in partnership with State Mental Health Agencies in Maryland and New York as part of the NIMH-funded RAISE initiative (1-3) and enrolled 65 adolescents and young adults with early psychosis suggestive of schizophrenia across two sites (Baltimore, MD, and New York, NY). Each team included a full-time Team Leader (licensed clinician), full-time Employment/Education Specialist, half-time Recovery Coach (licensed clinician), and 20% time Psychiatrist. Teams used assertive outreach strategies and shared decision-making to engage participants in care. Teams provided services for up to 2 years, using a critical time intervention model (8) with the goal of helping people stabilize their psychiatric conditions, reintegrate with school/work/family, and transition to appropriate community services and supports. Participants provided informed consent. Participating institutions' Institutional Review Boards approved study procedures.

What makes good fidelity measures?

Optimal fidelity measures are: informed by evidence, good proxies for intervention components being measured, objective, and drawn from readily-available information. Routine service logs or billing data support many fidelity measures (4). For example, such data have been used to document whether assertive community treatment teams are, indeed, delivering services intensively and whether clients are being served by multiple staff (4). Other easily obtainable, objective, pre-existing data can address structural requirements (e.g., minimum staffing and after-hours coverage) and processes of care (e.g., presence of completed side-effect checklists serve as indicators that side-effects were assessed). Such measures may be most useful in determining whether an implementation is minimally adequate as opposed to, for example, discriminating among exemplary programs (5).

Even when extensive administrative data exist, some topics are best addressed by self-reports from service users (6). For example, clients' ratings of whether staff used shared decision-making would be preferable to asking staff.

An Example: Measuring fidelity to a team-based intervention for people with early psychosis

This report by the RAISE Connection Program's lead researchers and intervention developers expands upon implementation findings reported elsewhere (1), describes how we monitored treatment fidelity using measures based on the principles stated above, and provides fidelity findings.

RAISE Connection Program researchers worked with the lead developers of core treatment domains (team structure and functioning, psychopharmacology, skills building, working with families, supported employment/education) to determine performance expectations for each domain and the operationalization of those expectations using information commonly available electronically to programs that bill for services (hereafter, "administrative data"). [A table included in an online supplement to this column lists performance expectations for these program domains and their operationalization into fidelity measures.] We followed this approach to enhance generalizability even though, as a research project that wasn't allowed to bill for services when new and funded by Federal research dollars, we couldn't ourselves use claims data to extract data on service use. Rather, we relied on research staff to extract information from routine service logs maintained by clinical staff and from specific fields in medical records (e.g., medication records). All data came from such objective field as opposed to reading progress notes (1).

In addition to these measures derived from program data, researchers worked with treatment-domain leads to identify questions to ask clients to determine whether, from clients' perspectives, intervention components had been implemented. [A figure included in the online supplement lists the questions and presents data on clients' responses.] These questions were embedded in structured research interviews participants completed at 6-month intervals post-enrollment (1) and provided corroboration for fidelity measurements obtained from program data. For example, we measured the expectation that, "The psychiatrist and client regularly review medication effectiveness and side effects" both by noting psychiatrists' completion of standardized side-effect monitoring forms and by asking clients, "How much did your Connection Team psychiatrist bring up the topic of medication side effects?" [see table in the online supplement].

For many fidelity measures, RAISE Connection Program had no pre-existing standard to adopt for what constituted "acceptable" performance. Rather, data collected during the project were used to generate expectations based on actual performance.

Findings

Both teams met or exceeded most performance targets [see table in the online supplement for a summary of teams' fidelity results from program data, measured both across time and

for the project's final complete quarter]. Data from client interviews also indicated high fidelity to the treatment model. The large majority of clients reported that teams paid attention to their preferences about jobs and school, made treatment decisions—including medication decisions-- jointly, and responded quickly.

An example of how we reviewed fidelity is by using the expectation that teams see clients in the field as needed. [A figure in the online supplement illustrates the performance expectation, set by consensus of the intervention developers, that at least 10% of clients meet with team members in the community, excluding visits with Employment/Education Specialists.] As shown in the online figure, both teams exceeded the performance expectation. Such figures, especially when they show performance of multiple programs, are useful in spotting deviations from expectations and performance outliers, and also are clear reminders of program expectations, change in performance over time, and how one's own program compares with other programs.

We also used fidelity measures based on program data to examine differences between teams. Recovery Coaches had different styles of providing services across sites, with Recovery Coach in Site 1 providing almost all services in a group format and Recovery Coach in Site 2 providing a mix of group and individual sessions.

Fidelity is a team-level measure, yet many fidelity measures are composed of aggregated client-level data (e.g., whether a client has had an adequate trial of an antipsychotic). These measures can be used to generate exception reports (e.g., lists of clients for whom there has been no meeting with a family member) that can be fed back to teams and supervisors to identify areas for improvement. By study end, we were able to provide data to teams from such exception reports and share data with supervisors.

For some expectations (e.g., that Employment/Education Specialists accompany clients to school or work when clinically indicated and desired by the client), we expected only a small fraction (e.g., 10%) of clients to endorse the item because the service may be relevant for few clients. For such measures, small but non-zero findings provide proxy measures indicating that treatment components were implemented. For other treatment components, we expected most clients to endorse the item because the component (for example, shared decision-making) was relevant to all participants.

Measuring fidelity efficiently and using fidelity findings

A core challenge in measuring treatment fidelity is to do so reliably and without breaking the bank, ideally by using data already being collected for other purposes. While research studies may rely upon reviewing videotapes or site visits to observe program implementation, such efforts can be too laborious for broad implementation. Bringing model programs to scale calls for cost-effective, sustainable approaches to measuring fidelity. Increasingly, payers' contracts with programs have dollars at stake with respect to maintaining program fidelity. Fidelity data used for such purposes need to be reliable and objective, requirements that may not be met by data from summary impressions of site visitors or small samples of observations.

Routine service logs will support many fidelity measures so long as they note, for each contact, the client, staff involved, whether family were present, and location of service (e.g., office versus community). Presence of routine clinical forms, such as those included in the RAISE Connection Program Treatment Manual, both support the intervention and can be used to document that those intervention components occurred (7).

While obtaining fidelity data from claims data and other pre-existing sources minimizes data collection/compilation burden on staff, as a fallback to using administrative data to measure fidelity, payers may specify data that programs are required to submit, and those submissions can be verified, in toto or at random, via site visits. Designing, building, debugging, and implementing an accompanying chart abstraction system is cumbersome for short-term use but offers an alternative when abstraction from electronic claims isn't possible.

As data accrued, we were able to see that most expectations appeared reasonable (Table, online) and used early data from these teams to revise staffing and performance standards for new teams being rolled out in New York under the OnTrackNY initiative (9). For example, lower-than-expected rates of metabolic monitoring led to adding part-time nurses to OnTrackNY teams.

Even when data aren't sufficient to establish precise performance thresholds, such data let program managers identify outliers (e.g., a team that never provides services off-site). Such outliers needn't indicate poor performance, but they point to areas for further investigation and follow-up, perhaps via site visits. For program start up, knowing that the service exists may be sufficient. If stakeholders become concerned that clients who are in need of the service aren't getting it, then a more nuanced measure would be called for.

Site visits can be costly and time consuming for routine fidelity monitoring, particularly in large systems, but can be helpful to reinforce training and to elucidate factors underpinning unusually good or poor performance on fidelity measures derived from administrative data.

Programs, and their funders, need to budget for fidelity measurement as core program costs. Building such data-reporting requirements into contracts helps ensure adequate budgeting. Fiscal bonuses for meeting performance expectations help incentivize performance. As noted above, while we don't always know *a priori* what good performance looks like, often we can define what is minimally adequate, so we can specify, and monitor, accordingly. With such data, we can identify good outliers (e.g., teams with high rates of engagement) and feature them in efforts to improve performance.

Conclusions

Programs can use routinely available data to determine whether many key components of an intervention have been implemented. Fidelity data from multiple sources indicate that the RAISE Connection Program was implemented as intended, including program elements related to shared decision-making as well as the range of expected clinical interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part with Federal funds from the American Recovery and Reinvestment Act of 2009 and the National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN271200900020C, Lisa Dixon, Principal Investigator; by the New York State Office of Mental Health; and by the Maryland Mental Hygiene Administration, Department of Health and Mental Hygiene. The authors thank Robert Heinssen and Amy Goldstein at NIMH for their efforts bringing the RAISE initiative to fruition and thank Dianna Dragatsi, Jill RachBeisel, and Gayle Jordan Randolph for their ongoing support. Jeffrey Lieberman was the original Principal Investigator on the contract with NIMH, and the authors thank him for his foresight in assembling the original research team.

References

1. Author Citation Rfb. Under Review.
2. Author Citation Rfb.
3. Lieberman JA, Dixon LB, Goldman HH. Early detection and intervention in schizophrenia: a new therapeutic model. *JAMA : The Journal of The American Medical Association*. 2013; 310(7):689–90.
4. Essock SM, Kontos N. Implementing assertive community treatment teams. *Psychiatric Services*. 1994; 46(7):679–83. [PubMed: 7552558]
5. Wisdom JP, Knapik S, Holley MW, Van Bramer J, Sederer LI, Essock SM. Best practices: New York's outpatient mental health clinic licensing reform: using tracer methodology to improve service quality. *Psychiatric Services*. 2012; 63(5):418–20. [PubMed: 22549525]
6. Essock SM, Covell NH, Shear KM, Donahue SA, Felton CJ. Use of clients' self-reports to monitor Project Liberty clinicians' fidelity to a cognitive-behavioral intervention. *Psychiatric Services*. 2006; 57(9):1320–3. [PubMed: 16968764]
7. National Institute of Mental Health. [May, 2014] Coordinated Specialty Care for First Episode Psychosis. Manual II Implementation. Available at : <https://nimh.nih.gov/health/topics/schizophrenia/raise/csc-for-fep-manual-ii-implementation-manual.pdf>.
8. Herman D, Conover S, Felix A, Nakagawa A, Mills D. Critical Time Intervention: An empirically supported model for preventing homelessness in high risk groups. *The Journal of Primary Prevention*. 2007; 28(3-4):295–312. [PubMed: 17541827]
9. Center for Practice Innovations. [August 1, 2014] OnTrack NY. Available at <http://practiceinnovations.org/CPIInitiatives/OnTrackNY/tabid/202/Default.aspx>.