EARLY PSYCHOSIS INTERVENTION ONTARIO NETWORK

Implementing a volunteer peer fidelity assessment in Ontario Early Psychosis Intervention programs:

What did we learn?

A project of the Standards Implementation Steering Committee

September 2018

camh

The Early Psychosis Intervention Ontario Network (EPION) is a province-wide volunteer network of service providers, persons with lived experience, and families. EPION currently includes over 50 programs and satellite partners across Ontario. The network facilitates collaboration, training, resource sharing, and quality improvement efforts. EPION is funded by the Ministry of Health and Long-Term Care. For more information, visit http://help4psychosis.ca/ or http://eenet.ca/the-early-intervention-in-psychosis-for-youth-community-of-interest/.

PROJECT TEAM

RESEARCH TEAM:

Janet Durbin
Provincial System Support Program, Centre for Addiction and Mental Health

Avra Selick
Provincial System Support Program, Centre for Addiction and Mental Health

Gordon Langill
Canadian Mental Health Association, Haliburton, Kawartha, Pine Ridge Branch

Chi Cheng
Northern Ontario School of Medicine

Suzanne Archie
St. Joseph's Healthcare, Hamilton: McMaster University

Donald Addington
University of Calgary

STANDARDS IMPLEMENTATION STEERING COMMITTEE (SISC):

Gordon Langill, Chair
Canadian Mental Health Association, Haliburton, Kawartha, Pine Ridge Branch

Catherine Ford
Ontario Ministry of Health and Long-Term Care

Chi Cheng
Northern Ontario School of Medicine

Eleanor Baker
Family member, Toronto

Janet Durbin
Centre for Addiction and Mental Health

Jessica McConnell
Family member, Barrie

Julie Beaulac
The Ottawa Hospital

Karen O'Connor
Canadian Mental Health Association, Peel Branch

Kevin Barclay
Champlain Local Health Integration Network

Moshe Sakal
Life and Spiritual Coach, Toronto

Sarah Bromley
The Slaight Centre, Centre for Addiction and Mental Health

Shannel Butt
Canadian Mental Health Association, Elgin Branch

Suzanne Archie
St. Joseph's Healthcare Hamilton

Suzanne Robinson
Central West Local Health Integration Network

SUGGESTED CITATION

Standards Implementation Steering Committee. (2018). Implementing a volunteer peer fidelity assessment in Ontario Early Psychosis Intervention programs: What did we learn? Centre for Addiction and Mental Health and the Early Psychosis Intervention Ontario Network: Toronto, Ontario.

Table of Contents

MAIN MESSAGES

- Routine fidelity monitoring is a process that can support adherence to program standards, program quality improvement and sector improvement initiatives; however, fidelity assessments are not routinely conducted in Ontario community mental health programs. This project tested the feasibility and value of a volunteer peer fidelity assessment process for Ontario Early Psychosis Intervention (EPI) programs.

- Key components of the process included: three person assessor teams (2 volunteer assessors from EPI programs and an evaluator from CAMH); a validated First Episode Psychosis Services Fidelity Scale; and visits to the program to learn about EPI delivery. A central project team provided training and coaching to enhance consistency and accuracy of ratings. Nine programs from across Ontario received an assessment and 20 assessors were trained to conduct the assessments

- Many elements of the assessment process supported the consistency and accuracy of ratings. The team worked well – the EPI peer assessors had credibility with the programs and understood program practices, and the CAMH staff brought knowledge of program evaluation and of systematic data collection. The site visits were important to learn about the programs. The detailed manual, training and coaching provided by the central team guided all components of the process.

- The fidelity reviews were valued by the EPI programs. They appreciated feedback on strengths, challenges and improvement suggestions, which some have started to act on. The common assessment created a potential opportunity for programs to share practices and collaborate on improvement projects.

- The fidelity reviews showed that program performance was satisfactory or exemplary for many model elements, although there was variation. One area where programs tended to receive lower ratings was in the psychosocial treatments domain (e.g., CBT, weight management, substance use treatment). A common suggestion was for more consistent documentation of practice and increased use of manuals and protocols to standardize how and when services were provided.

- A number of feasibility issues emerged. The assessment process required considerable time from both the programs and the assessors. Assessors required training and an opportunity to gain experience with the process. Given the assessor turnover experience with this project, maintaining a pool of qualified assessors will likely be a challenge. Additionally, central supports are required to organize the assessments, delivery assessor training and aggregate fidelity results.

- There was value to using an existing validated tool. However, small refinements are needed to increase relevance in Ontario. Preliminary work to develop supplementary items specific to the Ontario EPI Program Standards is underway.

- Based on the positive feedback received, the EPI sector has already started another round of assessments, seeking to increase efficiencies while maintaining the value of this peer team model approach to fidelity assessment. Sustainability and spread would significantly benefit from formal commitments from health system stakeholders for funding and support.

The following next steps are recommended for the EPI sector:

1. Conduct a follow-up survey to assess the impact of the fidelity assessment on the quality of services in the participating programs.
2. Continue working towards routine sector-wide fidelity assessments in Ontario EPI programs.
3. Explore strategies to increase efficiencies and reduce the burden on assessors and programs (e.g., alternate in-person assessments with self-assessments, incorporate remote assessment strategies).
4. Explore opportunities to support the use of fidelity results to improve the quality of care, such as through quality improvement projects and engagement with system partners involved in quality improvement work.
5. Explore other funding strategies to support sustainability of routine fidelity assessments in EPI programs (e.g., fee for service, partner with other funding organizations to support cost).
6. Engage other sectors to learn from their experiences with implementing fidelity assessments.

MAIN REPORT

Introduction

Early Psychosis Intervention (EPI) is a team-based treatment model for young people experiencing psychosis, that was developed based on reviews of research evidence. The implementation of EPI programs in Ontario began in 1994 and has been advanced by an active network called the Early Psychosis Intervention Ontario Network (EPION). EPION has wide volunteer participation by EPI program staff and managers, as well as representation by persons with lived experience, family members, decision makers, and researchers.

From 2004 to 2007, the province released significant new funding for the development of EPI programs. While the capacity of programs to provide EPI increased, this occurred in the absence of provincial standards.[1] In 2011, Ontario's Ministry of Health and Long-Term Care published the *Early Psychosis Intervention Program Standards*[2] and, shortly after, it formed the Standards Implementation Steering Committee in collaboration with Early Psychosis Intervention Ontario Network (EPION). The Provincial System Support Program (PSSP) at the Centre for Addiction and Mental Health (CAMH) committed in-kind support to the Committee for planning, monitoring, evaluation, and other activities to help sector implementation of the Standards. The longer term aim of the Committee is to develop systematic objective processes for monitoring the delivery of services and supporting quality improvement.

As an initial step to learn about current practice in relation to the Standards, the Standards Implementation Steering Committee conducted two province-wide EPI program key informant surveys during 2012 and 2014. The surveys identified the creative and innovative work being done in EPI programs and also the areas where more support is needed to meet the Standards. Of note, key findings included the need for monitoring, evaluation, networking and accountability. This is where the fidelity assessments can contribute.

Fidelity refers to the extent to which delivery of an intervention adheres to the standards, guidelines, or protocol that characterize it. Fidelity scales are a tool for measuring fidelity, defining measurable criteria to judge whether a program is delivering an intervention as it was intended. Fidelity assessments can guide program improvement work and provide a common standard for assessing quality of services delivered across a sector. If applied routinely with Ontario's EPI programs, fidelity assessments can flag

[1] Cheng, C., Dewa, C.S., & Goering, P. (2011). Matryoshka Project: lessons learned about early intervention in psychosis programme development. *Early Intervention in Psychiatry*, 5, 64-69.
[2] Ministry of Health and Long-Term Care. (2011). Ontario Early Psychosis Intervention Program Standards. Toronto, ON. Retrieved from: http://www.health.gov.on.ca/english/providers/pub/mental/epi_program_standards.pdf

issues where a system response may be appropriate and can help to ensure clients seeking services at any EPI program across the province will receive the same, high quality care.

In 2016, a research team, led by Dr. Donald Addington, published an evidence-based, validated scale to measure EPI delivery — the First Episode Psychosis Services Fidelity Scale (FEPS-FS)[3]. The FEPS-FS created an opportunity to advance monitoring of EPI program delivery in Ontario. However, measuring fidelity can be challenging. Fidelity assessments traditionally involve site visits and are labour intensive. They draw on multiple program data sources, including administrative data, chart reviews, and staff and client interviews. Data from these sources are integrated using a consistent rating process outlined in a rating manual. Preparing assessors — who often have different backgrounds and levels of expertise — for this task is challenging. Further, while the FEPS-FS assesses what have been identified as the core components of EPI services, it does not address all of the Ontario Standards and expectations for program delivery.

The present pilot project was initiated to test the feasibility and value of using an innovative peer assessment method to assess the fidelity of selected Ontario EPI programs using the FEPS-FS. The Standards Implementation Steering Committee formed a project team to support the implementation and evaluation of the pilot. Peer assessor teams of volunteer EPI program staff and PSSP evaluators were formed to conduct the program reviews. The pilot was completed in nine programs.

The present report describes the review process and the results of the evaluation. Also included are recommended next steps related to fidelity monitoring in EPI programs.

Project Objectives

1. Develop and implement a peer fidelity assessment process with a subset of Ontario EPI programs.
2. Evaluate the assessment process, considering consistency, feasibility, and value to programs.
3. Review fidelity results, considering how they could be used to support program and sector improvement.
4. Make recommendations for future fidelity monitoring.

[3] Addington, D. E., Norman, R., Bond, G. R., Sale, T., Melton, R., McKenzie, E., & Wang, J. (2016). Development and testing of the first-episode psychosis services fidelity scale. *Psychiatric Services*, *67*(9), 1023-1025.

Fidelity Assessment Process

The project team developed and oversaw the implementation of the assessment process, which was comprised of five core components. These included: a peer assessor team model; use of the FEPS-FS scale to evaluate program fidelity to the EPI model; site visits to each participating program; a post-visit consensus rating meeting; and preparation of a structured final report. Each of these components is described in more detail below.

1. Peer assessor team model

The assessor teams included 3 members - 2 volunteer EPI program staff (who assessed a program different from the one they worked in) and 1 evaluator/implementation expert provided through in-kind support from the Provincial System Support Program at CAMH. The combination of EPI program expertise and evaluation expertise was intended to strengthen the credibility, accuracy, consistency and timeliness of the review. Additionally, the hope was that using volunteer EPI program staff as assessors would build capacity within the sector to conduct the reviews and create a sustainable review strategy.

The project team recruited the volunteer assessors through a general call to EPI programs and PSSP staff. Assessors attended an initial two-day, face-to-face training session with Dr. Addington, followed by phone meetings to review the assessment process and the support materials in more detail. The project team prepared all support materials and organized team meetings. The materials included a rating manual and tools, based on the original research study but expanded to ensure the process for data collection and making the ratings was clear and consistent.

An assessor team was formed for each program review based on assessor availability. Thus, assessors who conducted more than one assessment typically worked with a different team for each one. The team always included at least one EPI program staff and one PSSP staff.

2. First Episode Psychosis Services Fidelity Scale

Dr. Addington developed the FEPS-FS through a structured knowledge synthesis process that included systematic reviews, international expert consensus, and pilot testing in two countries.[4] The scale includes 31 items which address team function, assessment processes, medication management practices, provision of psychosocial supports, and service access and continuity. Each item is rated on a five-point scale from 1 ("Not implemented") to 5 ("Fully implemented"). Rating criteria are item specific and are intended to measure both quality and reach (e.g., number of clients receiving). A score of 4 is considered satisfactory performance and 5 is considered exemplary adherence. *See Appendix A for the full scale.*

[4] Addington et al. 2016

3. Program Site Visit

Each assessment required a two-day, in-person visit by the assessor team to obtain first-hand information. The data sources and collection processes were planned in advance and included:

a. A structured review of 10 randomly-selected client charts
b. Interviews with program staff, clients, and family members
c. Observation of a team meeting
d. Review of administrative data (e.g., number of new referrals over the last 12 months, number of currently enrolled clients) and other program material (e.g., policies, training materials, public education materials, brochures)

Interview guides, data extraction templates, and a detailed agenda were provided to guide each visit and enhance consistency. The programs provided administrative data and policy documents in advance, obtained any required approvals (i.e., for external staff to talk to program clients and families and review client charts), and scheduled the interviews. Assessors had the option of providing high-level feedback to the program at the end of the second day.

4. Consensus rating meeting

After each visit, the assessor team attended a phone meeting with Dr. Donald Addington and the project team to systematically review what was learned about program delivery for each scale item and reach consensus on a final fidelity rating. By having Dr. Addington's input at every meeting, we hoped to enhance rating consistency across teams while building assessors' knowledge and skills. Feedback from the meetings was used to refine the manual and provide more clear and specific rating guidance.

5. Structured final report

Each program received a final report written by the assessment team that followed a structured template developed by the project team to simplify report writing and make them consistent across teams. It included an overall summary of program strengths and improvement opportunities as well as item results, including the fidelity rating, evidence to support the rating, relevant contextual information, quality improvement recommendations and additional comments. The narrative reinforced positive practices and provided guidance for improvement.

Implementation Supports

The project team managed the implementation process and guided assessors and EPI program staff throughout the review process. The aim was to increase consistency and, as possible, minimize the burden of effort to assessors and EPI staff. Supports included:

- **Assessor training and support**: Assessors received a two-day, in-person training session, a virtual half-day follow-up training session, and ongoing access to a central support person for any questions arising throughout the project.
- **Site support:** Each EPI program received step-by-step instructions on how to prepare for their site visit, including the materials they needed to provide to the assessment team as well as the meetings and interviews to schedule. Sites were also assisted to meet organizational privacy and ethics requirements including the client approvals to conduct the review.
- **Tailored manual and tools:** Assessors received a detailed manual that included definitions and scoring instructions for each scale item, the site visit protocol, and data collection tools, including interview guides, meeting observation guide, and a client record data abstraction template.

The project team included the Standards Implementation Steering Committee chair, a PSSP scientist, a dedicated project coordinator and a small amount of administrative support. The coordinator assumed overall responsibility for implementing the above components, including developing the tailored materials, and served as a central resource to answer questions and problem solve until final reports were submitted. The coordinator also formed the assessor teams for each review, scheduled site visits, and organized follow-up consensus meetings. The administrative assistant booked travel and accommodations for assessors.

The pilot project consisted of two phases, each including four or five pilot sites, to allow for ongoing feedback and improvements to the process.

Evaluation

Objectives

The project evaluation aimed to assess the following:

1. The effectiveness of the process for promoting rating accuracy and consistency
2. The feasibility of the process
3. The value of the assessments
4. The relevance of fidelity results for program and sector quality improvement work.

Method

Feedback was collected from multiple sources (see Table 1), including:

- Assessors, using focus groups and time tracking sheets
- Program staff, using interviews
- Consensus meeting observation
- Fidelity reports, using ratings and comments

The focus groups and interviews, conducted at the end of the project, were audio recorded and transcribed.

The data analysis was guided using questions that the project team developed based on the evaluation objectives. Qualitative data (from the focus groups and interviews, consensus meetings) was coded and summarized into themes related to the questions. Data from time tracking sheets were summarized by assessor type (EPI or PSSP staff) and by task. The fidelity ratings (total scale score, domain score, item score) and narrative comments in the fidelity reports were summarized across programs to identify commonalities in current practice, delivery challenges, and improvement opportunities.

Early findings were reviewed with participants for validation purposes. Ethics approval for the evaluation was received from the Research Ethics Boards at CAMH, Lakehead University, and St Joseph's Care Group.

Table 1: Data collection

Data source	Method	Participants
Assessors	Assessor focus groups after project completion	3 focus groups, total of 12 assessors
Assessors	Time tracking log completed after each assessment in a structured template to record time spend on each component of the project.	25 time records completed for 9 assessments
Program staff	Semi-structured interviews with staff from participating programs after project completion	5 interviews (representing 6 programs*)
Consensus rating meeting	Observation of each consensus rating meeting	9 consensus rating meetings
Fidelity reports + written program response	Completed after each fidelity review by assessors using a structured template, optional written response by site appended to report	9 final fidelity reports

*one staff manages 2 of the assessed programs

RESULTS

Participation

Program participation

Five programs participated in the phase 1 evaluation in February and March 2017. Feedback collected helped refine the process for the four programs in phase 2 of the project, which occurred from May to July 2017. The programs were located across the province and ranged in size and delivery model (see Table 2).

Table 2: Pilot fidelity review participating programs

Program	Rurality	# delivery sites	Clinical full-time equivalents	Host organization	Region
1	Mixed	1	1	community	North
2	Mixed	4	10.7	community	West
3	Mixed	1	5	hospital	West
4	Urban	1	14	community	Central
5	Urban	1	6.8	hospital	Central
6	Mixed	3	8.5	hospital	East
7	Mixed	2	6	hospital	East
8	Mixed	1	1.8	hospital	East
9	Urban	1	13.1	hospital	East

Assessor participation

In total, 15 EPI staff, from nine programs, and five PSSP staff, from five regional offices received training and committed to conducting reviews. Of these, six left before the end of the project due to job changes or lack of time. Sixteen assessors completed at least one and up to four assessments (two on average).

Effectiveness of the process

This section summarizes feedback on the effectiveness of the fidelity assessment process for promoting rating accuracy and consistency, in relation to each of the core components.

Assessor team model

Both the assessors and program staff felt that the assessor team model, with the combined PSSP and EPI staff, provided a good mix of complementary skills for making accurate and consistent ratings. As front-line clinicians, program staff on the assessment team brought their clinical knowledge of EPI care to the assessment process, which gave them understanding of the context and practices. They also had

credibility and were able to build rapport with staff in the programs they were assessing. The PSSP staff brought knowledge of program evaluation and of systematic data collection and rating processes. While there was a steep learning curve, pairing new with experienced assessors was a suggested strategy to build capacity. Teams generally functioned well but could benefit from more clearly defined roles, including an identified lead.

Neither the programs nor the assessors reported discomfort with the experience of assessing a peer or being assessed by a peer.

> ***"I felt we were really well balanced. We called ourselves the A Team. We worked really, really, really well together."***

First Episode Psychosis Services Fidelity Scale

Both the assessors and program staff were generally positive about the FEPS-FS. They felt that the scale addressed the key elements of the EPI model and provided a structured approach for making assessments. Using a validated, standardized scale gave weight to the assessment, and having an objective scale increased the comfort of peer assessors in rating other programs.

Most scale items were felt to accurately assess the quality of care. For a small number of items, however, concerns were raised that a lower score might not be an accurate reflection of the care being provided. For example, ratings for some medication-related items were based on best practices for schizophrenia while Ontario programs also accept clients with affective psychosis, which has a different treatment protocol and, thus, different quality criteria. Additionally, the rating criterion for a comprehensive assessment was considered by some to be too strict.

> ***"It was a good scale but it wasn't totally in line with EPI Ontario standards."***

In general, participants felt that scale items and ratings aligned with quality of care. They suggested that item-specific challenges could be addressed in the final report by explaining how each item was defined and the research supporting it, and by instructing programs to focus more on the comments than on the score for items that do not align with the Ontario context. However, this issue may require more attention if fidelity monitoring is spread to additional programs that are less comfortable with disregarding low scores, or if results are used to compare programs or to show sector performance.

Program site visit

Overall, respondents felt that the assessment protocol, including the in-person site visit and multiple data sources, contributed to rating accuracy.

Assessors reported that being on site helped them ensure their ratings were accurate and develop rapport with program staff. Staff interviews were a rich data source that gave them a fulsome understanding of how the program operated. Although clients and family members did not always remember specific details about their treatment, these interviews gave assessors a feel for the overall culture of the program and whether it was meeting client needs. Materials were helpful when provided by the programs ahead of time, and it was challenging when they were not provided.

The most challenging data source was the chart review. There was wide variation in the structure and completeness of client charts, and finding the information needed to make the ratings was often difficult and time consuming. In the absence of chart documentation, assessors had to rate activities as if they were not occurring, and there was concern that some ratings did not accurately reflect program practice. Quality and consistency of documentation was an important improvement area flagged for most programs. Due to this challenge, assessors varied in their approach to the chart review, which compromised rating consistency across teams.

> ***"I did two assessments, and there [were] probably 20 different kinds of charts. So some were electronic, some were a mix; and even within an agency, every chart wasn't the same."***

An additional concern was the size of the sample for the chart review and, specifically, whether a review of 10 charts could provide an accurate representation of practice in larger programs, such as those with over 200 clients. Strategies suggested to address this issue include reviewing more charts in larger programs or giving more weight to interview feedback.

The only data source the assessors did not think was necessary was the team meeting observation. The assessors felt they could get information about team function through interview feedback and that observing the team meeting was not an efficient use of the limited time onsite.

Consensus meeting and rating process

The consensus meeting constituted an important tool to educate new assessors on the nuances of scoring. Even after they gained experience, situations occurred where the rating was unclear and it was helpful to have external guidance to make the final ratings. The assessors felt that the consensus meeting improved the consistency of ratings across teams and recommended that it remain a part of the model.

> ***"[The consensus meeting] was great for clarification, for affirming and for setting us straight if we were wandering off the beaten path."***

Summary of effectiveness results

Overall, assessors felt that the fidelity assessment process enhanced consistency, and program staff generally reported that the ratings were accurate. However, some challenges were noted that should be reviewed if the process is to be expanded. Client charts were of variable quality, leading to some inconsistencies in how ratings were determined. Additionally, a larger sample of charts may provide a more accurate view of practice in larger programs. While the FEPS-FS fidelity scale was generally seen as a useful tool, additional work is needed for a small number of items to ensure they are relevant and meaningful for Ontario programs.

Although the assessment protocol needs some refinement, it worked for all types of programs in the pilot project, including those that were multi-site or part of a larger network. Still, the process for assessing programs in different network arrangements needs more consideration, as decisions on when to assess sites separately or as a single program were not always obvious.

Feasibility of the process

Assessors

The assessment required a substantial time commitment, more than the participants or project leads had anticipated (see Table 3 for a list of resources needed per assessment). The actual time spent ranged widely across assessors, but on average the PSSP assessors contributed 66 hours and the EPI assessors contributed 45 hours per assessment. This difference may have occurred in part because PSSP staff took on lead roles in managing the process and facilitating report preparation. The overall average was 53 hours per assessment.

"I think it's a very heavy piece of work to take on… I underestimated that."

The assessors spent a considerable amount of time working on each final report. Particularly challenging was completing the report after the site visit since the assessors did not have time protected for this task back at work. This delayed completion of the final report for some programs. Suggestions to manage this problem included reducing content requirements, pre-populating sections, where possible, identifying clear roles and responsibilities for team members, blocking time for report writing in advance, and ensuring that assessors and their managers have realistic expectations of the time required so their workload can be properly managed.

All but one of the reviews was completed by a three-person team, which was the preferred number of assessors to ensure all the site data collection activities could be completed in a timely way. However, it was a continual challenge for the project coordinator to negotiate schedules and last-minute substitutions were required on multiple occasions. Also, 30% of the originally trained assessors dropped

out during the project. For this model to be sustained, ongoing recruitment and training of new assessors would be required.

Despite the time needed, assessors felt that this work is important and that it would be feasible to donate staff time for this purpose a few times a year, at least for larger programs. However, this volunteer sample may not reflect wider views in the sector.

Programs

Preparing for a fidelity visit was a considerable amount of work that generally required about two to three days of staff time (see Table 3). A time-consuming task was compiling the advanced materials for assessors. Reducing the number of items needed could lessen the burden. Additionally, this task will be easier if sites put in place more consistent tracking of required data, such as the number of public education sessions.

The site visit was also time-consuming to host. However, program staff noted that having assessors on site helped staff understand the importance of the assessment and increased the likelihood that the results would be used. Overall, despite the time needed to prepare, the feedback was that it was feasible and important to participate. Program staff also felt that preparation for future assessments would be easier now that they better understand what materials are needed and have more time to prepare.

> ***"If we're asking our clients to do (these sorts of) assessments and put in effort for recovery, we have to be willing to do the same thing… We can't just say we're doing a great job when we have no evidence that we're doing a great job."***

One unexpected challenge for programs was managing the ethical and privacy concerns related to allowing external individuals to talk to clients and review client records. Most agencies do not have experience with this type of work and the approval process was often unclear and time consuming to sort out. Based on the pilot experience, the project team developed a consent form template and project description that may speed up the process in future assessments. However, the experience highlights the variation in practices for projects that are considered program or quality improvement/ evaluation rather than research.

At some sites the preparation process was also challenged by low staff engagement. While most sites described staff as enthusiastic about the project, some were apprehensive or disinterested. Site leads emphasized the importance of telling program staff about the upcoming assessment well in advance, framing it as an exciting opportunity to improve client care and showcase their work, and emphasizing that it is not about evaluating any individual's work.

"I kind of sold it as - this is an awesome opportunity. We hear anecdotally from the hospital, from consulting psychiatrists, primary care, that, you know, you guys have an awesome service. Okay, great. But let's actually dig in and see what we're doing."

Project team support

During the pilot project, the project team tested and refined the assessment protocol and tools (rating manual, data collection and reporting tools). With this developmental work completed, future assessments will likely require less support from the project team. There will still be a need for a coordinator with expertise in fidelity reviews, to organize and coordinate site visits, provide ongoing support to the assessors and the sites, and conduct the rating consensus meeting. New assessors will also need to be recruited, trained and supported to address turnover, and experienced assessors will need ongoing coaching and training refreshers. Continued administrative support will be needed for scheduling and booking travel.

Resource requirements and cost

While the actual resources needed for each assessment varied widely, Table 3 provides estimates of the average resources used to in this pilot project.

Table 3: Average resource use per assessment

Support	**Time/resources** (average per assessment)
Assessor time • *Includes preparing for the assessment, 2-day site visit, making the ratings, and preparing the report*	• PSSP assessor = 66 hours average • EPI assessor = 45 hours average • Total = 159 hours (53 hours per assessor)
Site time • *Includes preparing for the assessment, and supporting the assessors during the 2 day site visit.*	• 2-3 days of staff time prior to site visit to prepare documents for assessors. • 2 days to host site visit
Travel/ accommodation costs	• ~$1500 per assessment (based on 2 nights' accommodation & travel but varies widely)
Administrative support time • *Includes booking travel/accommodation and managing the reimbursement process*	• 6 hours per assessment
Coordinator time • *Includes protocol/tool development, ongoing support, forming assessor teams, evaluation.*	• 0.4-0.8 FTE during pilot (varied depending on stage of project)*

* This includes time for the evaluation and will be lower when process becomes more operational.

Summary of feasibility results

While the assessment process required considerable time from both the programs being reviewed and the assessors, the consistent message from participants was that this is important work and is feasible if programs are willing to make the reviews a priority. The extent to which these views are shared more

widely in the sector is unknown. Central coordination and support resources were also required, and are significant.

Assessors and program staff discussed the question of assessment frequency. While both groups felt that reassessments are important, they thought that programs need sufficient time to implement changes. They suggested reassessments every two to three years might be feasible to sustain while allowing programs to implement changes.

The need for on-site visits also was discussed, given that less labour intensive options are available. Self-assessments, for example, could complement the external/on-site assessment (e.g., self-assessment at 18 months, external assessment at three years), though participants agreed that they would not want to entirely replace the formal assessments with self-assessments.

Value of the process

> ***"I would say in the grand scheme of my experience in community and mental health programs, this is a real jewel in the crown."***

Assessors

The assessors enjoyed participating and expressed interest in continuing in the future. They valued seeing how other teams operate, learning about different ways that EPI services can be delivered, and bringing ideas back to their teams. Assessors also found it validating to see that others are struggling with the same challenges.

> ***"I think the benefit is actually going out and seeing how another site works …because a lot of time they're doing great things."***

Programs

Program staff indicated that the fidelity assessment was valuable to identify areas for improvement, support advocacy for resources, and validate program strengths. While the results typically were not surprising, having issues formally identified sparked a discussion on how to address them. Even some staff who were originally apprehensive, found the process to be a helpful opportunity to discuss program delivery, celebrate what they were doing well, and be honest about where they could do better. Sites with internal evaluation capacity still found it valuable to get an external perspective on areas for improvement and suggested strategies. The assessment also helped teams see the importance of collecting data and how it can be used to improve care.

Programs also saw the fidelity assessments as having value for the sector, with potential to support more consistent, high-quality EPI services across the province and make it easier for clients who are transferring programs. Sector-wide fidelity assessments could be used to identify higher performing programs and better practices that could be shared. Finally, the fidelity assessment provided an impetus for programs to refer back to the Ontario EPI Standards as a guide to program delivery.

> ***"I think it would be very important for every service to go through this, because I would be concerned that we say we meet the standards when we -- none of us actually meet the standards."***

Typically, managers reviewed the report with staff and some programs also shared the report with others, including their directors, advisory boards, and regional or network bodies. Many programs reported plans to use the fidelity results for program improvement and some have started using the results to inform a quality improvement plan. Some programs also described new documentation practices they put in place as a result of the fidelity assessment.

> ***"I know we've already used our report in things like hospital accreditation to say what we're doing for QI processes and that kind of thing. With our advisory board, we're already using the report to prove what we're doing in our program, so it's helpful."***

Some programs suggested it would be helpful if they received more support on how to use the fidelity results. This could include review the report, identifying priorities for improvement, and developing a quality improvement plan. There was concern that, without support, even programs with the best intentions might not actively use their fidelity results. Another suggestion was to bring programs together to share learnings and work together to plan improvement projects.

The main concern raised was how long programs had to wait to receive their final reports (generally took two to four months, but in some cases up to seven months). Program staff said they would like processes put in place to ensure reports are received in a more timely way.

Summary of value results

Both assessors and programs found the fidelity assessment to be valuable and, despite the time demand, most participants said they would be willing to participate again in the future. Routine fidelity assessments have the potential to be a powerful tool to support both sector and program improvement efforts, though programs may require additional support to utilize the results.

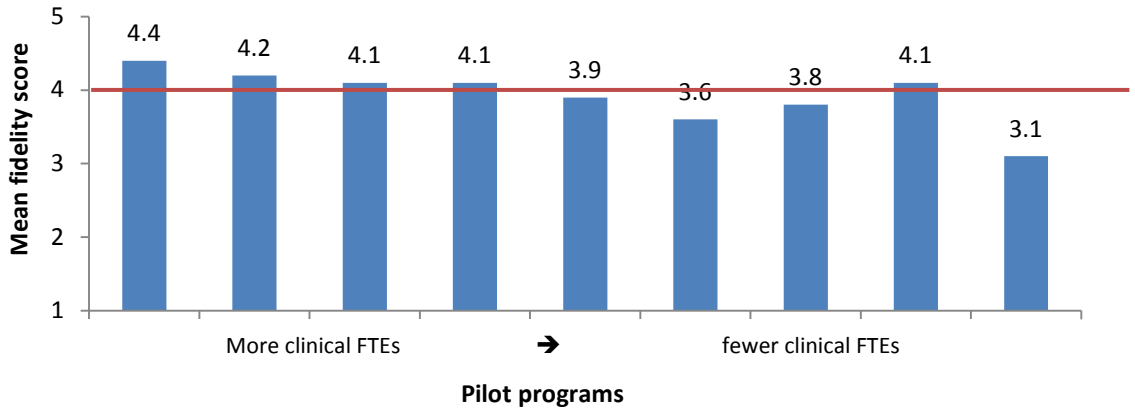
"People have been excited about this, have wanted this. And so, there is an appetite for people understanding- How well are we adhering to the standards? How can we improve things? And the iron is hot. This is the time."

Fidelity results

This section demonstrates what is possible to learn about EPI delivery from the fidelity assessments that can inform program and sector improvement work. The results represent the experience of the 9 participating programs and not necessarily EPI program practice in Ontario more broadly.

Figure 1 reports the mean overall fidelity score for each program in the pilot. When considering how closely a program follows the EPI model, a rating of 4 indicates that the program follows the EPI model closely. While most program scores were close to 4, the scale was effective in distinguishing between different levels of fidelity. This variation creates an opportunity to examine contextual factors that may influence practice, such as program resources, access to specialized staff, and catchment area size. In Figure 1, mean overall scores are ordered by program FTE size. Similar to our earlier sector surveys,[5,6] results suggest that smaller programs have more difficulty achieving a higher score. A larger sample of programs could explore this and other system-wide patterns that may be worth discussing or following up.

Figure 1: Mean fidelity scores across pilot programs



[5] Standards Implementation Steering Committee. (2012). *Implementation of Early Psychosis Intervention Program Standards in Ontario: Results from a Provincial Survey*. Centre for Addiction and Mental Health and the Early Psychosis Intervention Ontario Network: Toronto, Ontario. Retrieved from: http://eenet.ca/sites/default/files/pdfs/EPI-Program-Survey-Final-Report-October-2012-pdf-pdf.pdf

[6] Standards Implementation Steering Committee. (2015). *After Release of the Ontario Early Psychosis Intervention (EPI) Program Standards: Results of the 2014 EPI program survey of current practices in relation to the Standards.* Centre for Addiction and Mental Health and the Early Psychosis Intervention Ontario Network: Toronto, Ontario. Retrieved from: http://eenet.ca/sites/default/files/pdfs/EPION-SISC-Survey-2-Report_July-31-2015-FINAL_3.pdf

Figure 2 groups the 31 fidelity items into five domains that align with different areas of program practice. These include:

- Access and continuity, such as timely access, proactive outreach, crisis support, and communication with inpatient services
- Team practice, such as multidisciplinary team, weekly meeting, and psychiatrist role on the team
- Assessment and care planning, such as comprehensive initial assessment, patient and family involvement, and annual reassessment
- Pharmacotherapy, for example, medication prescribing
- Psychosocial treatments, for example, psychoeducation, supported employment, and psychotherapies, such as cognitive behaviour therapy (CBT)

The results show that programs were closer to the EPI model in domains related to access and continuity, team practice, and pharmacotherapy, and less so for the recovery domain. This finding is further illustrated in Figure 3, where mean scores for each scale item are reported. The wide range shows differences in extent to which programs are able to implement the model. Ratings are low for some psychosocial treatment items (i.e., CBT, weight management, substance use treatment) but also for a few team practice items (i.e., assigned a psychiatrist, team meetings). These items may be more challenging for smaller programs to implement.

Figure 2: Mean domain scores across pilot programs

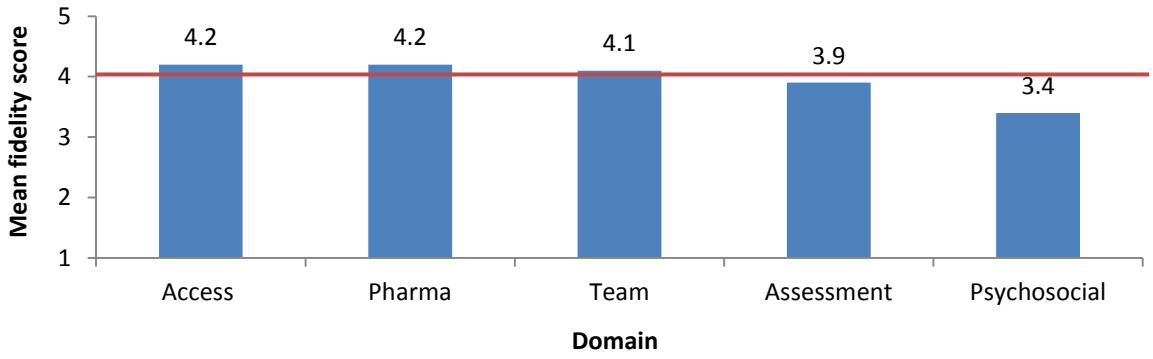
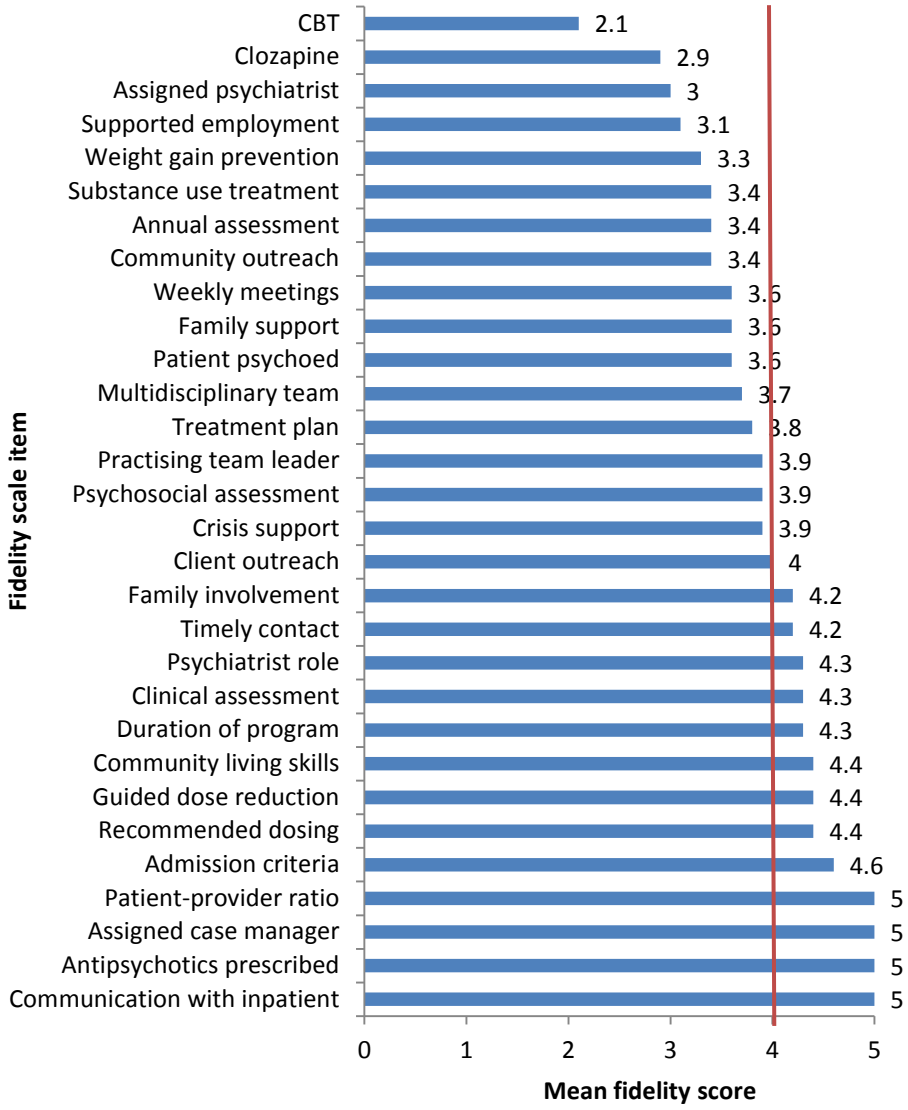


Figure 3: Mean item scores across pilot programs



The narrative comments in the reports highlight some of the common delivery challenges underlying these results, along with recommendations for improvement. Three are summarized below.

Standardization in delivery: A common pattern was that services were provided without using a standardized approach, creating a risk for inconsistency in delivering care. To address this, assessors suggested using manuals and protocols to standardize how and when services are provided.

Program capacity: For some care components, staff training and practices did not fully meet the EPI standards. One example was CBT. Many programs have staff with some CBT training and provide CBT-informed care, but fewer have certified CBT staff and offer a formal multi-session version of this intervention, as is required to achieve high fidelity. Supported employment is another area where access to specialist staff was lacking. Suggested improvement strategies for these items include accreditation

training for staff, partnership arrangements for referral to specialist providers, and clear delivery protocols.

Inconsistent monitoring and documentation: Programs varied widely in the number and type of tools they were using to monitor care delivery and, typically, in whether and when these tools were used. This variance contributed to inconsistent documentation of the care being delivered. Assessors suggested that programs streamline the number of tools they use and develop clear protocols for their use. Given its presence in many programs and in the Ontario system, a number of assessors suggested more consistent use of the Ontario Common Assessment of Need (OCAN). Regular use of this assessment could enhance fidelity on items related to psychosocial assessment, annual reassessment, and involvement of clients in the development of treatment plans. Also suggested was a tool to document treatment goals and progress. Importantly programs generally needed clear protocols so tools would be regularly used to document care.

The reports also offered details on how EPI programs deliver specific services, such as weight management (see Table 4).

Table 4: Fidelity report feedback on weight gain prevention (item 13)

Mean score = 3.3 across the 9 programs; range: 1-5

Implementation challenges:

- Most programs monitor client weight but do not systematically provide weight management support
- Without consistent documentation, it is difficult to ascertain what is actually being monitored and how often
- Even when weight management programs are offered, uptake of group activities by clients varies

Current strategies:

- Health education and healthy eating integrated into cooking group
- Peer coaches engaging with clients to encourage physical activities
- Parent education to support healthy lifestyles in clients

QI suggestions:

- Develop core content for structured, evidence-based weight management sessions
- Develop protocols for more consistent delivery and documentation
- Use local resources

Overall, the reports provided rich information about how programs are delivering EPI and how they can implement specific components of the model. It was suggested that creating an inventory of practices based on the reports and on the literature could help with quality improvement work across the sector.

Summary of assessment results

The fidelity results are based on a sample of nine programs and do not necessarily represent broader practice in the sector. However, the results align with our prior surveys of the sector,[7] which generally showed higher adherence for the more structural items of care (e.g., access and continuity, team practice, and pharmacotherapy domains) but lower adherence for psychosocial supports and therapies (e.g., recovery domain). The results also highlight innovative practices and common challenges. Combined, this feedback could inform quality improvement efforts within programs and across the sector. If administered routinely, fidelity assessments could also provide a systematic approach to monitoring the effects of improvement efforts.

[7] Standards Implementation Steering Committee, 2012; Standards Implementation Steering Committee, 2015.

Summary

This project tested a fidelity assessment process for Ontario EPI programs that included volunteer peer assessors, a validated fidelity scale, site visits to learn about the programs, and support from a central project team to enhance consistency and accuracy of ratings. The peer assessors and program staff both gave positive feedback about the assessment experience.

The assessors reported feeling well prepared and supported to conduct the reviews by the project team. The site visits both enhanced the accuracy of their reviews and provided an opportunity to learn about other programs. They felt that the team model, which combined PSSP and EPI staff, provided a good mix of complementary skills and helped make ratings more accurate and consistent.

Fidelity ratings showed that assessments can capture variations in the way programs adhere to the core elements of the EPI model of care. The programs reported that it was validating and encouraging for their strengths to be highlighted, and that identification of challenges informed follow-up conversations. Many program staff said they have already started improvement work based on the feedback, although several indicated they would welcome assistance. Program staff also felt the fidelity assessments have value for the sector - to share practices, work together on improvement projects, and strengthen consistency and quality.

The fidelity scale items remind program staff about the international and Ontario EPI standards that should guide their practice. The Ontario standards were never intended to be static and this pilot project may spark discussion about the current standards and need for refinements to align with new evidence about the EPI model and with the Ontario policy context. More broadly, this work contributes to building a culture of using data to guide improvement.

While the fidelity assessment process was generally effective, feedback highlighted some areas where the process could be improved. One was the chart review, which may benefit from a different sampling approach, clearer criteria for assessing practice, and possibly a screening process to determine readiness for a fidelity assessment. Additionally, while the fidelity scale mostly worked well, there were a small number of items that do not align with the Ontario context.

Regarding feasibility, the volunteer peer assessor model worked well but its sustainability is uncertain. The time commitment was considerable, and training and experience were both important components to ensure assessors had the necessary skills. However, there was high turnover and maintaining an ongoing pool of trained, available assessors will likely be a challenge. While the assessors felt that this is an important and feasible process, this was a self-selected sample and the extent to which their perspective is held by others in the sector is unknown. Continued implementation will also need the support of a project coordinator to conduct assessor training, monitor and improve the fidelity process, and manage the implementation logistics (e.g., scheduling reviews, forming assessor teams, preparing

sites). If fidelity results are to be aggregated for sector improvement work, this also would need to be resourced.

Based on the positive results of this pilot and high levels of sector interest, a second wave of assessments has already been planned. Refinements to the tools and process based on feedback from this pilot are underway, including preliminary work to develop supplementary fidelity scale items specific to the Ontario EPI Program Standards. It is worth noting that in preparation for this next wave of assessments, less than two years after the original training, only seven of the original 20 assessors are still available to participate.

It is a challenge beyond Ontario and beyond the EPI sector to develop fidelity review processes that are valid and feasible to sustain and spread. Options being explored in our jurisdictions include program self-assessments, remote telephone assessments, and technical support centers with dedicated trained assessor staff.[8,9] It will be important to learn more about these different options, considering feasibility and value for Ontario.

Fidelity assessments are ultimately intended to guide service improvement efforts. Across the system they offer a common platform for describing current practice, identifying quality gaps, and building capacity. We are currently exploring options to support EPI programs with this work. One is to establish a community of practice for EPI service providers. Another is to collaborate with provincial improvement initiatives, such as Health Quality Ontario and E-QIP (the *Excellence through Quality Improvement Project*, an initiative led by Addictions and Mental Health Ontario and the Canadian Mental Health Association, Ontario). Discussions about working together on sector improvement work are already underway, as are conversations with the Local Health Integration Networks. As a developing provincial data source, sustainability and spread would significantly benefit from formal commitments from health system stakeholders for funding and support.

[8] Rollins, A. L., McGrew, J. H., Kukla, M., McGuire, A. B., Flanagan, M. E., Hunt, M. G., … Salyers, M. P. (2016) Comparison of Assertive Community Treatment fidelity assessment methods: reliability and validity. *Administration and Policy in Mental Health and Mental Health Services Research*, 43, 157–167

[9] Margolies, P. J., Humensky J. L., Chiang, I-C., Covell, N. H., Broadway-Wilson, K., Gregory, R., … Dixon, L. B. (2017) Is there a role for fidelity self-assessment in the Individual Placement and Support model of Supported Employment? *Psychiatric Services*, 68(9), 975-978

Recommendations

Based on the results of this initial fidelity pilot, we recommend the following next steps for the EPI sector:

1. Conduct a follow-up survey to assess the impact of the fidelity assessment on the quality of services in the participating programs.
2. Continue working towards routine sector-wide fidelity assessments in Ontario EPI programs.
3. Explore strategies to increase efficiencies and reduce the burden on assessors and programs (e.g., alternate in-person assessments with self-assessments, incorporate remote assessment strategies).
4. Explore opportunities to support the use of fidelity results to improve the quality of care, such as through quality improvement projects and engagement with system partners involved in quality improvement work.
5. Explore other funding strategies to support sustainability of routine fidelity assessments in EPI programs (e.g., fee for service, partner with other funding organizations to support cost).
6. Engage other sectors to learn from their experiences with implementing fidelity assessments.

Acknowledgements

This work could not have been possible without the active participation and collaboration of many EPI program stakeholders. We want to thank everyone involved for their time, conscientious work, and thoughtful feedback. EPION has made this project a priority and consistently supported the work of the Standards Implementation Steering Committee. PSSP provided valuable in-kind assessor support and the PSSP assessors strengthened the consistency and quality of the process. Finally, the Standards Implementation Steering Committee provided valuable feedback and has enthusiastically championed this work. We look forward to continuing to build structures and processes to support high quality EPI care in Ontario.

APPENDIX A: First Episode Psychosis Services Fidelity Scale (FEPS-FS)