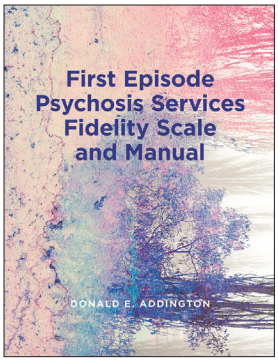
UNIVERSITY OF CALGARY
Press



FIRST EPISODE PSYCHOSIS SERVICES FIDELITY SCALE AND MANUAL

by Donald E. Addington

ISBN 978-1-77385-209-6

THIS BOOK IS AN OPEN ACCESS E-BOOK. It is an electronic version of a book that can be purchased in physical form through any bookseller or on-line retailer, or from our distributors. Please support this open access publication by requesting that your university purchase a print copy of this book, or by purchasing a copy yourself. If you have any questions, please contact us at ucpress@ucalgary.ca

Cover Art: The artwork on the cover of this book is not open access and falls under traditional copyright provisions; it cannot be reproduced in any way without written permission of the artists and their agents. The cover can be displayed as a complete cover image for the purposes of publicizing this work, but the artwork cannot be extracted from the context of the cover of this specific work without breaching the artist's copyright.

UNIVERSITY OF CALGARY
Press
press.ucalgary.ca

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

GENERAL GUIDELINES

Background

The First Episode Psychosis Services Fidelity Scale (FEPS-FS-1.0) is used to assess the degree to which mental health teams deliver specialized care comprised of a range of evidence-based practices to people experiencing a first episode psychosis and their families. Program fidelity refers to the extent to which delivery of an intervention adheres to the protocol of an evidence-based program model. Fidelity scales provide a list of objective criteria by which a program is judged to adhere to a reference standard for the intervention. The scale can be conceptualized as an outcome measure for implementation research or as a quality measure for assessing structure and process indicators in health care (Donabedian, 1966).

The First Episode Psychosis Services Fidelity Scale (FEPS-FS-1.0)

The FEPS-FS was developed using a standardized methodology for developing fidelity scales (Bond et al., 2000). The first stage involved three steps: (1) a systematic review of randomized controlled trials of first episode psychosis teams to identify the components of successful programs, (2) an assessment of the level of evidence supporting those components, and (3) a Delphi consensus process using international experts to identify the essential components (Addington, Mckenzie, Norman, Wang, & Bond, 2013). The second stage included three steps: (1) Developing descriptions of the core components, (2) defining behaviorally anchored criteria for ratings on a 5-point continuum for each component, and (3) development of a rating manual that described detailed procedures for data collection and scoring. The scale was tested for face validity, interrater reliability, and feasibility on programs in Canada and the United States . Results of the testing showed good interrater reliability and face validity. Compared with three other scales published (the Early Assessment and Support Alliance (EASA) scale, the RAISE Connection fidelity tool, and the EDEN scale, developed for Evaluating the Development and Impact of Early Intervention Services in England), the FEPS-FS has fewer components but the highest proportion of components common to all these scales (D. E. Addington et al., 2016). The design of the scale and the evidence base upon which it was developed permits assessment of programs that are based on different models and operating in different health systems.

The FEPS-FS-1.0 is a modified version of the original scale. Modifications were made in the course of and following two multi-center fidelity studies. The first, conducted in Canada, was a study of 9 programs in Ontario using an onsite review by trained clinicians, actively involved in health care delivery and health care evaluators (Durbin et al., 2019). The second study was a representative longitudinal cohort study of 36 US programs using a central remote assessment team and phone interview. The reliability of 33 items of the revised scale were shown to be good to excellent (Addington, Noel, Landers, & Bond, 2020). After the studies were completed, minor editorial changes were made to improve the wording of components and rating

criteria. Two new components reflecting components missed in the original scale were added to form the 35 components of the FEPS-FS 1.0. The two additional items included item 10, The age range served by the program, and item 35, Attention to fidelity. These are clear and concrete items that do not affect the reliability of the FEPS-FS 1.0 compared with the FEPS-FS revised version.

The scale is available in two forms: a 35-item Team scale and a 20-item individual patient scale.

1. The team scale can be rated in one of three ways:
    1.1. Site visit when expert reviewers visit a site.
    1.2. Remote assessment. Data is collected at the site from health record review and administrative data and interviews are conducted from a central site.
    1.3. Self Report. The site collects the data, interviews staff, and reports the results.
2. The individual scale is focused on the services offered to an individual patient. The scale assesses the service components received by the patient (Appendix G). This version of the scale was developed to assess and quantify the quality of the care received by individual patients who did not receive care from a team-based service. It was originally designed for use in a randomized controlled study comparing services offered in standard care versus specialized team care. This version of the scale has not been tested in research studies.

The scale was used as a self-report measure in a national study in Italy (Addington et al., 2020). Instructions on how to use the FEPS-FS 1.0 as a self-report measure are included in Appendix H. The self-report version does not differ from the observer-assessed version of the scale. The reliability of using the measure as a self-report measure has not been tested. Training in use of the scale is recommended for using the scale as a self-report measure just as it is for use by an external assessor.

The terminology used in mental health services varies across countries, service delivery organizations, and the professional groups delivering the services. The purpose of the scale is to assess the quality of services received, not to focus on terminology. For example, the scale and the manual refer consistently to patients, but not clients. Fidelity raters should use the terminology of the programs and services that they are evaluating. Thus, if the program or service uses the term clients rather than patients, the rater should use the term clients. Professionals delivering the services are also described using different terms. The term case manager means different things in different programs. In one program, the case manager does outreach and connects the patient to community services. In another program the case manager is the person primarily involved in coordinating the care within the team and perhaps for delivering a specific component of the services. It is a task of the fidelity rater to identify the services delivered by the different professional groups within their program and to use the program's own terminology to identify the various staff.

Different programs serve different patient diagnostic groups. This is a challenge for the fidelity rater, the fidelity sponsor, and the program. The research on team-based care for people with a first episode psychosis has focused on those with a first episode of a schizophrenia spectrum disorder. However, some programs also serve those with an attenuated psychosis syndrome (also known as the clinical high risk or ultra high risk groups), bipolar disorder, and/or a major depressive disorder with psychotic features. These different groups require different treatments, and the fidelity scale can be used in a way that adjusts for these different patient groups. These adjustments are defined as general principles and specific components for specific

disorders. It is up to the funders of specific fidelity reviews to determine the goals and objectives of the review. The most common focus for the fidelity reviews is for patients with a first episode of a schizophrenia spectrum disorder not including the attenuated risk syndrome. If that is the situation, the health records should be selected to reflect this population.

General Principles:

- Reliable initial diagnosis.
- Record numbers in groups.
- Review health records for population of interest.
- Exclude patients not in population of interest.
- Examine each group with sufficient numbers independently.

If there is a reason to assess fidelity for other groups, it is important to establish that there are sufficient numbers in those groups. To establish fidelity for these groups some components need to be dropped, as shown in the following table.

PATIENT GROUP	ITEMS TO OMIT
Recent Onset (0–5) years Schizophrenia Spectrum disorders	None
Bipolar	19 Initial antipsychotic 20 Antipsychotic dosing 21 Clozapine
Attenuated Psychosis Syndrome	13 Early Intervention 19 Initial antipsychotic 20 Antipsychotic dosing 21 Clozapine

The FEPS-FS is not designed to be a substitute for more detailed operating guidelines including country-specific or health-system-specific guidelines, such as provincial or state guidelines. The scale has been designed to focus on evidence-based services that are specific to or adapted to individuals with a first episode psychosis. Therefore, important general health system policies such as those addressing cultural adaptation, legal requirements for privacy, notification of dangerousness, or requirements for involuntary treatment are not addressed.

FEPS-FS-1.0 Manual

The *First Episode Psychosis Services Fidelity Scale Manual* provides a guide for scoring the FEPS-FS-1.0 and is designed to increase the reliability (consistency) of ratings across different sites and assessors.

The Manual provides the following:

- A definition and rationale for each component in the Fidelity Scale.
- A list of data sources to inform the ratings for each component.
- Decision rules to help score each component correctly.
- Site interview data-collecting tools: interview guides, health record abstraction guide.
- Site fidelity assessment preparation guide.

Overview of the Scale

The FEPS-FS 1.0 (see Appendix F) contains 35 program-specific components. Each component on the scale is rated on a 5-point scale ranging from 1 to 5. In the language of implementation this ranges from 1, "Not implemented" to 5, "Fully implemented". In the language of quality this ranges from 1, "Poor Quality" to 5, "Excellent Quality". In both languages, a 4 means good or satisfactory. The standards used to establish the anchors for the "fully implemented" ratings were determined through a variety of expert sources. The scale assesses the services received by program patients, the training received by the care providers, and how the team works together to engage and retain patients and deliver coordinated and evidence-based care.

Services Delivered and Staff Roles

Staffing patterns, professional designations, and individual roles vary significantly from one organization to another, and across health care jurisdictions and countries. To address this, the FEPS-FS-1.0 ratings focus on assessing the services received by the patients, rather than professional designation of the person who delivers the service. In practice, assessors need to adapt questions to fit with the staffing pattern of the program being reviewed. For example, Cognitive Behavioral Therapy (CBT) might be delivered by a psychologist or by a trained counsellor who is a social worker by profession. Case management may be provided by a mental health professional who is called a case manager or care coordinator, or by someone called a counsellor or recovery coach who may have additional roles such as CBT or individual resiliency training. The scale also assesses the professional training of the staff and the specific training received in order to fulfil their role on the team and deliver the services provided to patients.

Training Fidelity Assessors

Trained assessors should conduct the fidelity assessments. Training varies depending on the experience and knowledge of the fidelity rater. It usually involves a two-day training program followed by two teleconferences with the trainer to discuss consensus ratings of real programs. The training can be delivered in-person or remotely.

The first day of training addresses the following:

- The scale development process
- Review of individual components
- Review of the Manual
- Review of the process used for a fidelity assessment
- Review of best practices for implementing fidelity measures and other health care quality indicators

The second day's training involves case-based training.

After conducting their first few fidelity assessments, raters should have the opportunity to review the process and the ratings with the trainer.

Fidelity Assessment Process

The scale can be used for onsite fidelity reviews, remote fidelity assessment, and self-assessment. The scale was originally designed and first evaluated using an expert in-person site interview method. It has since been adapted and tested for remote assessments including both remote data collection and staff interviews. The fidelity scale has also been used as a self-report measure. The scale and the recommended sources of data are the same in each data collection method.

Preparing for a Site Fidelity Assessment

Fidelity assessments require advanced preparation by all participants to ensure that assessors have time to speak with different program stakeholders and receive the information necessary to make the ratings.

Assessor Role:

- Review the fidelity Manual, scale, and data collection tools in advance of the site assessment.
- Review any documents sent by the site in advance (see data source #1 in Data Collection and Data Sources section below for more detail).
- Complete any required training on ethical evaluation practices and health information privacy regulations.
- Communicate with the central coordinating unit to schedule the assessment date and organize the phone-in schedule.
- Ensure all necessary paperwork related to privacy/confidentiality is completed. (Requirements will vary from program to program.)
- Note: The assessors should meet in advance of the site assessment by teleconference to introduce themselves, discuss initial impressions from advanced materials, and confirm roles/approach for interviews. This can be conducted by phone.
- Arrange a specific time to train the team leader and the health record abstractor.

- Interviews for a single program should be scheduled for one day and should start with the interview with the program manager.

Program Role:

- Identify the lead contact with the fidelity review team.
- Create a schedule for site interview, line up staff interviews, pull active patient health records, and send assessors any important documents/reports. (For full instructions refer to Appendix A: Preparing for a Fidelity Assessment.)
- Ensure all necessary internal approvals (i.e., administrative, ethics) are in place and required paperwork has been completed by the assessors.
- Identify the person responsible for health record abstraction.
- Communicate with the central team to schedule the assessment date.

Central Team Role:

- Liaise with the site and assessors to organize and schedule the site visits.
- Support ethics and privacy processes.

The most successful fidelity assessments are those in which there is a shared goal among the assessors and the program site to understand how the program is progressing and delivering evidence-based practices.

Data Collection and Data Sources

A detailed guide for programs undergoing a fidelity review can be found in: Appendix A: Preparing for a Fidelity Assessment.
The assessor team will need to review three data sources: existing documents and administrative data sent in advance, data abstracted from the health records of active patients, and interviews with staff and patients. A schedule for the interviews or site visit will be prepared in advance by the site to ensure the evaluation runs smoothly

1. Existing documents and administrative data

Documents including policies, practices, detailed program description (including all program components), education materials, and routine reports/admin data (e.g., staff FTEs, admission and discharge statistics, etc.) should be provided by the program in advance. The full list of documents to include can be found in **Appendix A: Preparing for a Fidelity Assessment**. These documents need to be reviewed before the site assessment.

NOTE: *Only aggregate, de-identified data should be shared in advance.*

2. Data abstraction from active health records

a. **Ten active, randomly selected health records of patients who have spent one year in the program.** The health record data abstraction will require approximately 4 hours. The central team will work with each site to develop a randomization process. The records should be selected from patients who have been receiving services for at least one year to ensure patients

have had adequate time to receive the services. Where programs are new or so small that they do not have 10 patients who have received services for a year, the assessment of the health records needs to be reconsidered in light of the purpose of the fidelity review. If the purpose of the study is to strictly compare programs against a standard which requires a size of service that has a meaningful impact on community services, the small program can be rated according to the proportion of 10 records that meet fidelity criteria. If the purpose of the fidelity review is to check on the processes of care as a new program is being established, the rating can be calculated based on the proportion of health records of patients who have received care for one year. The health record review can be undertaken by a local abstractor who will require brief training from the team responsible for the fidelity review and an orientation to the health record by a clinical team member of the first episode psychosis service. More detail on selecting patient health records can be found in **Appendix A: Preparing for a Fidelity Assessment**.

b. **Health records of the last five patients who have had a hospital admission after joining the program.** The focus of this health record abstraction is to identify the data required to rate Component 31, *Communication Between FEPS and Inpatient Service,* and Component 32, *Timely Contact after Hospital Discharge*. Where programs are new or so small that they do not have five patients who have been hospitalized, the assessment of the health records needs to be reconsidered in light of the purpose of the fidelity review.

Any program requirements to support privacy or access from both an ethical and logistical perspective should also be confirmed ahead of time. The **Health Record Review Checklist** (Appendix B) can be used to extract the relevant data from the health records. It is designed to be completed without the use of any information that could be used for personal identification.

3. Interviews with staff

A range of program staff should be interviewed during the fidelity assessment. At all programs it is important to interview the program manager/team leader, one case manager/care coordinator, the psychiatrist or other prescriber, and the supported employment specialist. These interviews will be organized in advance by the program, and interviews are conducted individually. The specific configuration of interviews with staff will depend on the team composition and functions. The assessor should interview all staff needed to obtain the necessary data to complete fidelity ratings.

See Appendix C for the **Fidelity Interview Guide**. It is helpful to review the interview schedule with the site lead at the beginning of the first day of the site visit to clarify the roles of each staff member who will be interviewed.

It is the responsibility of the assessors to ensure that, when required, informed consent is received at the beginning of each interview and to make clear that participation is voluntary and that a decision to not participate will not affect staff member's employment in the program. Site-specific consent forms will be provided to each assessor team, and the signed consent forms will remain onsite and stored by the program. Detailed notes should be kept during the interviews to support the final fidelity ratings.

NOTE: *No identifying information (e.g., names of staff, patients, or families) should be included in assessor notes. Notes also should not include any comments on individual work performance.*

Confidentiality and Data Storage

It is important to ensure that proper confidentiality protocols are in place for each site interview and that any data collected is stored in an appropriate manner. Prior to or at the beginning of the site review, all assessors must sign the necessary confidentiality forms. These vary depending on the data collected during the site review and may include notes from the patient, family and staff interviews, notes from the staff meeting, and the completed Health Record Review Checklist (Appendix B), as well as any documents or reports shared by the site. The information collected should not include any names or references to individual patients, family members, or staff or personal health information. No health records or identifiable information should leave the site.

The information will be kept in a secure location by the assessor (if paper-based, in a locked cabinet/office, and if electronic, in a password-protected file) until the final fidelity report has been produced. At that point, the assessors will send the files to an appropriate site storage depending on the purpose and regulations governing that purpose. If the assessment is part of a research project, the storage follows research requirements. If it is a quality improvement project, then local storage requirements should be in alignment with local procedures. Before sending the documents, assessors should complete a final review of the documents to ensure no individuals are identified.

How to Rate Components and Triangulate Across Data Sources

It is the task of the assessors to review and synthesize all data collected to determine the score for each component on the scale and to complete the final report.

How to rate components:

- Ratings should be made based on the scale as it is written. Any concern that the rating does not accurately reflect program practice should be captured in the comments section of the final report.
- The scale ratings are based on current behavior and activities, not planned or intended behavior.
- For multi-site programs, if service delivery differs across program sites, consider rating the sites separately or rate according to the higher performing program and describe the discrepancy in the comments section.
- If a period of time is not specified, then the rating can be based on service delivered at any point during the period of care.

Which data source to use:
The next section of this Manual lists all relevant data sources for each component. All the listed data sources for each component can be used to complete the comments section of the report, but we have included instructions on *which data source is the suggested primary source to identify the rating for that component*.

- **Wherever possible, program administrative data or health records should be used to determine the final component rating.**

- In general, if care related to a component is typically documented in the health record and is provided to all patients, then the health record is used as the data source. In this case, if the component is not documented in an individual health record, we assume it did not happen. It is of course possible that it did occur and simply was not documented. This possibility can be discussed in the comments but should not impact the rating.
- However, if a component is **not** routinely documented in the health record, or is provided only to a minority of patients, an alternative data source (e.g., staff interviews) can be used to support the rating. If this is a possibility, interviews will be listed as the appropriate data source for that component. For example, if cognitive behavioral therapy is not documented in the health record, the record cannot be used as a data source.
- **All** data sources may provide important contextual information that should be included in the comments section of the final report (though it may not necessarily impact the rating). For example, if the rating for a component is low and the policy review indicates that no procedures are in place to support the component, this may be a practice improvement area to flag. On the other hand, if the rating is low and the processes seem appropriate, it may be a documentation issue, and this may be flagged.

After Your Fidelity Review

If possible, time should be set aside after the interviews and after the review of the documentation to review data to explore discrepancies. Queries can be addressed with a follow-up email or call to the site. It is critical that the interview rating and documentation be reviewed while the information is still fresh.

After the site interview *a consensus rating meeting* may be scheduled with the assessors and a fidelity expert if available. The assessors present their ratings and rationale, obtain feedback from the expert, discuss differences, and agree on a final consensus rating. The fidelity expert will ensure scoring decisions are in-line with the intended use of the scale.

The assessors can prepare the fidelity ratings and/or report and send to the program or research team within an agreed upon period. Two weeks after the consensus rating meeting is a reasonable time frame. A **Feedback Report Template** can be found in Appendix H. The report can include:

- A high-level overview of findings, highlighting program strengths and opportunities for improvement
- Component fidelity scores
- Data sources used to determine each score
- Any contradictions between data sources
- Any additional relevant contextual information that might explain the score (e.g., problems with outreach in remote regions)
- Specific additional information requested per component (clearly specified in template)
- Additional assessor comments (e.g., Was this component difficult to rate? Do you feel the rating is valuable/ reflects program practice?)