

9. Program Evaluation

In this chapter we will briefly review the GAIN's integrated and progressive approach to assessment (section 9.1), how it is related to key facets of program evaluation (section 9.2), and specific issues in implementing the GAIN (section 9.3). The latter includes a list of specific tools we provide to help use the GAIN for program evaluation.

9.1 An Integrated and Progressive Approach to Assessment

As noted in chapter 1, the GAIN was designed to integrate research and clinical assessment in order to create an efficient process that maximizes both reliability and validity (see Dennis, 1998). It is less a single measure than a related set of measurement batteries that are designed to be part of a progressive approach to assessment. This approach includes:

- **Identifying who needs to be fully assessed** – The focus here is on relative brevity and simplicity of administration. This can be attained with a GAIN-Quick or another comparable measure (see Dennis, 2002 for a list of other measures).
- **Identifying participants for targeted referral** – The goal of assessing for targeted referral is to determine who needs crisis or brief intervention (e.g., by SAP, doctor) versus more detailed assessment and specialized treatment/referral. Decision rules about where to send a participant may be more complex (e.g., substance abuse, mental health, both). For this kind of screening, the assessment is more likely to use the GAIN-Q and possibly supplement it to support a brief intervention.
- **Comprehensive biopsychosocial** – The GAIN proper is used to identify common problems and how they are interrelated among people who are highly likely to be admitted to treatment. It requires more skill in administration and even more in interpretation.
- **Specialized assessment** – The biopsychosocial may identify areas where additional assessment by a specialist (e.g., psychiatrist, school counselor) may be needed to rule out a diagnosis or develop a treatment plan or individual education plan. The GAIN is designed to facilitate referral to this next level by using DSM-IV and collecting questions that specialists need to know.
- **Program-level assessment** – The combination of screeners, the GAIN and other data can be combined to provide an important tool for program management, evaluation and development. Specifically it can be used to identify unmet needs and provide feedback on impact of new approaches.
- **Evaluation research** – The same data can be used to support local program evaluations or research studies and combined into a cross-site database for comparison with other programs to examine less common subgroups and answer methodological questions.

In terms of evaluation research, the GAIN incorporates several methods checks, has been mapped onto numerous major epidemiological databases, and contains a number of items to support benefit cost analysis of substance abuse treatment. In addition to conducting methodological work on the GAIN, we also encourage people to compare it with other measures,

records and clinical judgment. User-friendly reports to help clinicians and participants interpret the results have been developed and are available for use.

9.2 Relationship of the GAIN to Program Evaluation in General

Stakeholders. In most program evaluations and treatment service research studies, there are multiple audiences, each with different needs, interests, priorities and timelines. It is essential to consult with all potential stakeholders and end users of research from the earliest planning stages. In fact, explicitly identifying the mutual self-interests of the various collaborators in the research endeavor is crucial to its success. To the extent that these often disparate needs and issues can be integrated, the project will be easier to implement and the resulting report more useful. In contrast, failure to consider the input of research partners will likely lead to roadblocks, if not outright subversion. Many evaluators consult with program or agency leadership, but often leave out other key groups such as the program's line staff members and participants or other closely related agencies (e.g., major referral sources). These groups can be involved via interviews, advisory board meetings, surveys, focus groups, participant observation, or full ethnographic studies (Dennis, Fetterman, & Sechrest, 1994). The GAIN helps address these issues by supporting multiple definitions and collecting information that is needed for different common stakeholder subgroups.

Timelines. Decision makers often require preliminary feedback on study results to guide their planning, particularly when the research is intended to explicitly inform practice or policy. Evaluators can often design the study to produce information that would be useful if provided in a timely manner, but may be of little or no use if provided years later. Examples include a needs assessment component, preliminary study data to support a funding request, preliminary evidence suggesting that an experiment is very harmful (or more beneficial than expected), the results of an individual's clinical assessment, or input on a staffing problem. Local evaluations can be particularly informative by providing quarterly management reports summarizing participant demographics, clinical characteristics and outcomes for the current period and to date. Where feasible, it is also desirable to compare performance against large contemporary datasets (see Guess & Tuchfeld, 1979 for one of the best examples of the U.S. government facilitating this comparison). The GAIN is set up to provide immediate feedback for clinicians and rapid cleaning/feedback at the program level for planners. Because the data is pooled across sites, it can also provide access to a wide range of comparison groups and facilitate dissemination of findings.

Logistical constraints. Unlike general clinical research, most program evaluations and treatment services research studies deal with ongoing entities and active operations. Therefore, formative evaluations can be useful to understand and improve the treatment program. Furthermore, although a summative evaluation on a particular procedure or service may be meaningful, it is often less useful to do this on an entire modality, program, or agency. For example, although early drug treatment evaluations focused on the effectiveness of the main treatment modalities such as outpatient, methadone, short-term residential, long-term residential, subsequent analyses have found considerable variation within modalities that severely limits the usefulness and generalizability of the original findings (Condelli & Hubbard, 1994). In fact, with the movement toward providing a continuum of care along which participants move, it is becoming more and more difficult to statistically separate the effect of a specific modality from total treatment

received. Considerable work has been done with the GAIN to allow it to be used across levels of care, age and a wide array of clinical subgroups (see section 1.3 and Dennis, Lennox, Funk & McDermeit, under review).

Questions. While every research study has its own specific questions, the list below mentions some of the most common questions (Dennis, Fetterman, & Sechrest, 1994):

- Who is being served?
- Who is not being served?
- What are the major participant subgroups?
- What services did they receive?
- What needed services did they not receive?
- To what extent are services being targeted?
- Are some service protocols more effective than others?
- Are some service protocols more cost-effective or benefit-cost efficient than others?

It perhaps overstates the obvious to note that the specific research question being asked determines the design of the study needed to answer it. Yet this simple equation is frequently neglected. Current literature often seems to overlook the fact that randomized designs, commonly viewed as the pinnacle method of basic clinical research, can actually only answer questions dealing with relative effectiveness and cost-effectiveness and are not very powerful in answering other types of questions. Failure to address the other questions with appropriate designs has often called into question the validity of field experiments (Dennis, 1990; 1994). Many existing measures fail to document the clinical severity of the population or services provided; others focus more on lifetime epidemiology and are not set up well to measure change over time or economic costs and benefits. The GAIN attempts to include questions to answer each of these questions in a balanced approach.

Other Evaluation Issues and Strategies. For a broader program evaluation, it may also be useful to review the recommendations Dennis, Perl, Huebner, and McLellan (2000) offered for improving the state-of-the-art of substance abuse treatment evaluation:

1. Identifying in advance all stakeholders and issues (reviewed above).
2. Developing conceptual models of intervention and context.
3. Identifying the population to whom the conclusions will be generalized.
4. Matching the research design to the question.
5. Conducting randomized experiments only when appropriate and necessary.
6. Balancing methodological and treatment concerns.
7. Prioritizing analysis plans and increasing design sensitivity.
8. Combining qualitative and quantitative methods.
9. Identifying the four basic types of measures needed.
10. Identifying and using standardized measures.
11. Carefully balancing measurement selection and modification.
12. Developing and evaluating modified and new measures when necessary.
13. Identifying and tracking major clinical subgroups.
14. Measuring and analyzing the actual pattern of services received.
15. Incorporating implementation checks into the design.
16. Incorporating baseline measures into the intervention.

17. Monitoring implementation and dosage as a form of quality assurance.
18. Developing procedures early to facilitate tracking and follow-up of study participants.
19. Using more appropriate representations of the actual experiment.
20. Using appropriate and sensitive standard deviation terms.
21. Partialing out variance due to design or known sources prior to estimating experimental effect sizes.
22. Using dimensional, interval and ratio measures to increase sensitivity to change.
23. Using path or structural equation models.
24. Integrating qualitative and quantitative analysis into reporting.
25. Using quasi-experiments, economic or organizational studies to answer other likely policy questions.

The GAIN was explicitly designed to support many of these approaches. The list of articles using the GAIN (at <http://www.chestnut.org/li/gain>) identifies several examples of it being used with different populations and analytic methods.

9.3 Planning to Implement and Use the GAIN

Exhibit 9-1 contains a list of common questions that need to be addressed when using the GAIN for clinical practice or research and is meant as a starting point, not an exhaustive list. But, these are issues better addressed sooner rather than later. In many cases Chestnut staff members work with a group of grantees (within a multi-site study) as a data-coordinating center to clean the data, generate management reports, create analytic files, and support cross-site analysis. Individual licensees typically manage their own site's data and analyses. We do have a policy of sharing our instruments, manuals, and syntax for minimal or no cost (with licensed users). In addition, the GCC can manage an individual licensee's data and analysis needs if arranged in advance (usually contractually). Please contact GAINInfo@chestnut.org if you are interested in these services.

Some of the specific tools available on the attached CD and our website (<http://www.chestnut.org/li/gain>) to support local program evaluation or treatment research are:

- Copies of the instruments (including an archive of prior versions).
- A crosswalk identifying what the items are used for and a core set for a shorter version.
- Maintained cross walks of data collected by version and site across over 100 projects to facilitate later cross-site or methodological analyses.
- Information on the HIPAA-compliant GAIN ABS online system for managing local versions of the instrument, generating individual level reports, transferring records and exporting (de-identified) data.
- SPSS syntax for cleaning data and creating scale scores for analysis and reporting.
- Crosswalks to existing standards (e.g., ASAM, DSM, GPRA, JCAHO, TEDS).
- Individual level reports to support specific interventions (e.g., Sampl & Kadden's 2001 MET/CBT5), motivational interviewing, interpretation and referrals/recommendations in general.
- Detailed psychometric norms by age and level of care.
- An electronic encyclopedia (more than 1,000 pages) describing calculated scales and variables used for analysis or clinical reports. Information is provided about each scale's

and variable's measurement type, description, calculation, interpretative ranges, and related references. (See Appendix H for a full listing of the scales and variables.).

Moreover, we continually add to our website and will continue to expand this list. It is particularly useful to check on the scales and variables file as it is being updated quarterly to include the most recent publications and changes to any syntax to address problems that have been identified by other users.

Exhibit 9-1. Common Implementation Questions

Design

1. Which instruments will you be using (e.g., GQ, GAIN-I, TxSI, M90, GCI, GCF, others)?
2. How will you be using the information (e.g., diagnosis, placement, treatment planning, outcome monitoring, program development/evaluation, research)?
3. From where will you recruit people?
4. What is the process to get the participant from identification to admission?
5. Do you need a screener or process of determining initial eligibility?
6. How will you implement and document random assignment or placement decisions that need to be made?
7. What safeguards are in place to protect the integrity of any randomization or quasi-experimental assignment decisions?
8. How will you document initial working alliance and the services provided?
9. When will you do follow-up (e.g., at 3, 6, 9, 12 months either post-intake or post-discharge). Will it be done in person or by telephone?
10. What will be your follow-up protocol (e.g., releases, locators, verification, reminders, trackers, monitoring, follow-up log)?

Clinical

11. What policies will you follow related to suicide risk or imminent danger to others, infectious diseases, victimization, and substance use during pregnancy?
12. Who will initially get and have access to the intake results? Are there implications for releases or business partnerships under HIPAA?
13. Who will initially get and have access to any follow-up data?
14. Who will have ongoing access to the records (based on role or ID) and who will monitor this process?
15. Have you identified critical issues or populations to be tracked for management reports (e.g., intake case flow, clinical characteristics, retention data, leading outcomes like urine tests or school or employment attendance)?
16. What steps will you take to get staff members to integrate the assessment into their actual practice?

Training and Supervision

17. What do you need in your study specific appendix (e.g., cross-site ID, local site ID, staff ID, time periods, instruments to be completed)?
18. Who will be responsible for training, supervising and certifying the staff members administering the GAINS?
19. How will you initially train people?

20. How will you train new people due to turnover?
21. Will you be granting continuing education credits for staff members?
22. What process of quality assurance (e.g., tapes, direct observation) and certification will be used locally?
23. Who will be the backup person when the trainer/supervisor is unavailable or in the event that she leaves?
24. How will you proactively identify and address staff members' concerns during the initial implementation phase?

Statistical

25. Who will be responsible for doing additional local reports or analyses?
26. Who will be involved in the data management or just receiving (de-identified) data from which to work?
27. Do the people responsible for #25 and #26 have SPSS or other software to support statistical analysis?
28. Who will provide syntax and information to help data managers use the data?

Logistical

29. Where will you do the assessments at intake and follow-up? Will they be private and secure? Will you have multiple remote locations?
30. Who will administer and (if applicable) score and interpret the GAIN? Will it be a few people or everyone?
31. Will you administer the GAIN online with GAIN ABS or use the paper version and data-enter responses after the fact?
32. Do the people involved with #31 have adequate equipment to do this?
33. Where will the hard and electronic data be stored?
34. Who will be responsible for installing and updating software?
35. Who will be responsible for managing the local database (and corresponding with data coordinating center where applicable)?
36. Do you have to accommodate people with special language, cultural or other needs?
37. Do you have to accommodate collaterals or children while doing the assessment?
38. Do you need any special releases for treatment or follow-up?