Fundamentals of Evaluation for Health Programme Planning

Dr Joseph D. Molina,
M.D., MSc (Public Health & Epidemiology)
Research Analyst, HSOR, NHG

Definition

“Evaluation involves assessing the strengths and weaknesses of programs... to improve their effectiveness.”
-American Evaluation Association

“Determines whether or not goals and objectives are met by assessing the level of conformance to evaluation criteria.”
- Clement and Wan, 2001

The Program

Evaluation attempts of any kind were acknowledged as critical but were conducted by only about half of the CCPs.

Forms/methods of the evaluations:
- anecdotal reports or feedback via newspapers, field reports, field visits, and responding agencies
- surveys often conducted by untrained outreach workers
- contracted evaluations with only two reporting useful findings

Response rates, when reported, were less than 20 percent

A few program directors mentioned resistance surrounding the idea of evaluation

Source: Elrod, Hamblen, Harris. Challenges in Implementing Disaster Mental Health Programs: State Program Directors’ Perspectives. The ANNALS of the American Academy of Political and Social Science 2006 604: 152

Why Evaluate

- Evaluation attempts of any kind were acknowledged as critical but were conducted by only about half of the CCPs.
- Forms/methods of the evaluations:
  - anecdotal reports or feedback via newspapers, field reports, field visits, and responding agencies
  - surveys often conducted by untrained outreach workers
  - contracted evaluations with only two reporting useful findings
- Response rates, when reported, were less than 20 percent
- A few program directors mentioned resistance surrounding the idea of evaluation

Source: Elrod, Hamblen, Harris. Challenges in Implementing Disaster Mental Health Programs: State Program Directors’ Perspectives. The ANNALS of the American Academy of Political and Social Science 2006 604: 152
**Framework for the Evaluation**

- **Review the program’s goal(s)**
- **Translate goal(s) to measurable indicators/outcomes**
- **Collect data from those enrolled and those not enrolled in the program**
- **Compare outcomes between those enrolled and those not enrolled**

---

**What should be evaluated**

Evaluation criteria – Donabedian’s triad:

a. **Structure** – quantitative information on:
   - Infrastructure of health services
   - The organisation
   - Types of services available
   - Resources available

b. **Process** – extent to which care/service provided meets the criteria (which may change over time)

c. **Impact** – level of health achieved as the result of the intervention

---

**The Logic Model**

**A Series of “If... Then” Statements**

- **If** you have access to funds and you can use them to accomplish your planned activities,
  - **Then** you can start working on your intended results.

---

**Resources/Input**

- **Your Planned Work**
- **Activities**
- **Output**
- **Outcome**
- **Impact**

Source: US CDC

---

**26 Sept 2013**

ASIAN COMMUNITY MENTAL HEALTH LEADERSHIP FORUM, SINGAPORE: 24 – 26 SEPTEMBER 2013
ASIAN COMMUNITY MENTAL HEALTH LEADERSHIP FORUM, SINGAPORE: 24 – 26 SEPTEMBER 2013

What should be evaluated

Outcomes

Health
- Clinical
- Health resource utilisation
- Patient-reported

Economic
- CBA
- CEA
- CMA
- CUA

Program objective 1:
To assess the coverage of Crisis Counseling Assistance and Training Program (CCP) for individuals affected by hurricane Katrina

Program objective 2:
To assess the effectiveness of Cognitive Behavior Therapy for Postdisaster Distress (CBT-PD) to address a range of cognitive, emotional, and behavioral reactions to disaster.

Program components:
- 8-10 sessions
- 1st 2 sessions on Psychoeducation
- Breathing Retraining and Behavioral Activation
- Program core (from session 3): Cognitive restructuring

Quality

Reach
- Review the program's goal(s)
- Translate goal(s) to measurable indicators/outcomes
- Collect data from those enrolled and those not enrolled in the program
- Compare outcomes between those enrolled and those not enrolled

Quality
- Effectiveness as reflected by intensity of post-traumatic stress at various time points
- Service volume over time
- Results of 12-item Short Post-Traumatic Stress Disorder Rating Interview
- Comparisons by state, gender, race/ethnicity

Reach
- Review the program's goal(s)
- Translate goal(s) to measurable indicators/outcomes
- Collect data from those enrolled and those not enrolled in the program
- Compare outcomes between those enrolled and those not enrolled
Evaluation designs and statistical methods are based on fundamental designs used in epidemiological research.

### How to Evaluate?

**Observational Studies**
- Cross-sectional
- Case-control
- Cohort
- Quasi-experimental

**Experimental Studies**
- Randomized controlled trials
- Quasi-experimental
- Cohort

### Evaluation Designs

1. **Before-after, program-no program design**
   - Compare outcomes within groups (before vs after), and between groups (program vs no program).

2. **Chronological series (single group time series)**
   - Multiple (serial) assessments done pre- and post-intervention.

3. **Before-after studies using national averages as comparison**
   - Multiple (serial) assessments done pre- and post-intervention.

### Graphs
- Quality of life indicators before and after intervention.
- Comparison of interventions with national samples.
How to Evaluate? – Evaluation Designs

4. Goal-based evaluation model
- Comparison is not with a parallel group
- Targets are pre-determined

5. Single group before-after (pre- post-) design
- Compare outcomes before establishment of the program with results after the program became available

Requirements:
- Baseline data should be available
- Similarity in the manner by which pre- and post-data are collected

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Useful in situations where there is no parallel control and when randomisation is not possible</td>
<td>Pre- &amp; post-data often not comparable due to quality/completeness</td>
</tr>
<tr>
<td></td>
<td>Maturation and historical effects</td>
</tr>
<tr>
<td></td>
<td>Lack of a parallel control limits interpretation of results</td>
</tr>
<tr>
<td></td>
<td>Results are often suggestive rather than conclusive</td>
</tr>
<tr>
<td></td>
<td>Regression to the mean</td>
</tr>
</tbody>
</table>

Advantage Disadvantages
- Pre- & post-data often not comparable due to quality/completeness
- Maturation and historical effects
- Lack of a parallel control limits interpretation of results
- Results are often suggestive rather than conclusive
- Regression to the mean

6. Post-test only design with nonequivalent groups
- When the evaluation is worked out after the program has began
- Assumes that no pretest observations are available
- Cannot rule out selection bias because of lack of baseline data
- Lack of time-relevant data precludes assessment of maturation effects
How to Evaluate? – Evaluation Designs

**7. One group post-test only design**
- No way to infer that treatment is related to any kind of change
- Assumes that no pretest observations are available
- Totally uninformative except to describe the “state” of the group

---

How to Evaluate? – Example 1

**Reach:** To assess the coverage of Crisis Counseling Assistance and Training Program (CCP) for individuals affected by hurricane Katrina

19 crisis counseling programs in 17 states

3 programs in disaster-declared areas
16 programs “undeclared” areas for displaced persons

Comparative findings:
- Programs showed excellent reach relative to external benchmarks provided by FEMA registrations for individual assistance and population characteristics
- The declared programs reached more people, but the undeclared programs provided more intensive services to fewer people with higher needs

---

How to Evaluate? – Example 2

**Quality:** To assess the effectiveness of Cognitive Behavior Therapy for Postdisaster Distress (CBT-PD) to address a range of cognitive, emotional, and behavioral reactions to disaster.

Referral → Pre-treatment → At 3rd session → At last session → 3 mos after last session

Review the program’s goal(s)
Translate goal(s) to measurable indicators/outcomes
Collect data from those enrolled and those not enrolled in the program
Compare outcomes between those enrolled and those not enrolled

---

How to Evaluate? – Example 1

**Reach:** To assess the coverage of Crisis Counseling Assistance and Training Program (CCP) for individuals affected by hurricane Katrina

Review the program’s goal(s)
Translate goal(s) to measurable indicators/outcomes
Collect data from those enrolled and those not enrolled in the program
Compare outcomes between those enrolled and those not enrolled

---
Quality: To assess the effectiveness of Cognitive Behavior Therapy for Postdisaster Distress (CBT-PD) to address a range of cognitive, emotional, and behavioral reactions to disaster.

Review the program's goal(s)

Translate goal(s) to measurable indicators/outcomes

Collect data from those enrolled and those not enrolled in the program

Compare outcomes between those enrolled and those not enrolled

Change in post-disaster distress

---

Suggested Readings


Rosen C, Young H, Norris F. On a road paved with good intentions, you still need a compass: Monitoring and evaluating disaster mental health services. In C. Ritchie, P. Watson, & M. Friedman (Eds.), Mental Health interventions following disasters or mass violence (206-223/2001). New York: Guilford Press.
