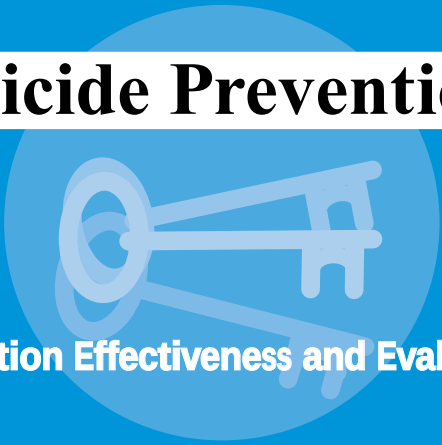
Lack of evaluation research is the single greatest obstacle to improving current efforts to prevent suicide.

– from an article in MMWR, April 22, 1994
Vol. 43/No. RR-6, page 6

Suicide Prevention



Prevention Effectiveness and Evaluation

National Center for Injury Prevention and Control

SPANUSA®
Suicide Prevention Action Network

Education Development Center, Inc.

Lack of evaluation research is the single greatest obstacle to improving current efforts to prevent suicide.

– from an article in MMWR, April 22, 1994
Vol. 43/No. RR-6, page 6

Suicide Prevention

Prevention Effectiveness and Evaluation

National Center for Injury Prevention and Control

SPANUSA
Suicide Prevention Action Network

Education Development Center, Inc.

This publication was supported by Grant #R13/CCR417100-01 from the Centers for Disease Control and Prevention (CDC). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC.

Suggested citation:
SPAN USA, Inc. (2001). *Suicide Prevention: Prevention Effectiveness and Evaluation*. SPAN USA, Washington, DC.

CONTENTS

FOREWORD

AFFIRMATION OF LIFE

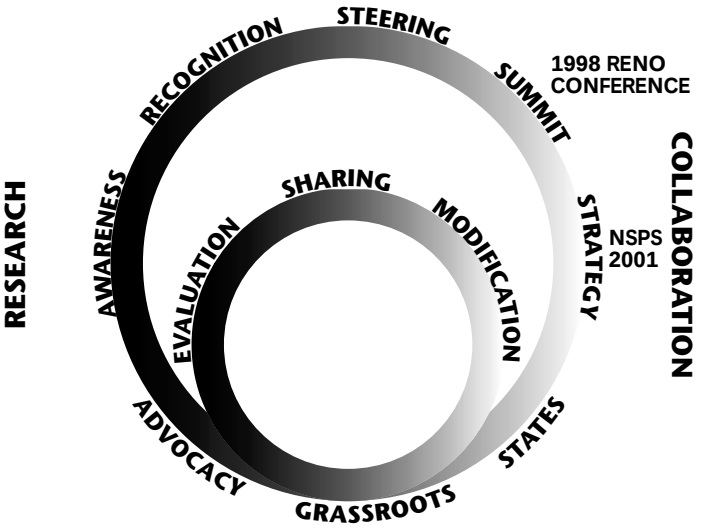


Figure 1. The SPAN Plan

BACKGROUND. For eight years following the suicide death of our daughter, Terri Ann Weyrauch, M.D., we volunteered in local and national organizations. We learned that even though there was a solid interest in non-government supported suicide research and professional continuing education, there was little dialogue among leaders of existing programs. Nor was there collaborative or consistent involvement in suicide prevention efforts. Additionally, few grassroots, non-professional survivors were involved or felt welcome to participate.

Lloyd Potter, Ph.D. (then with the CDC, Center for Injury Prevention and Control), was a U.S. representative to the 1993 World Health Organization Suicide Prevention Conference held in Calgary, Canada. Lloyd provided us with an advance draft of the paper that was later produced by the conference: **Prevention of suicide: Guidelines for the formulation and implementation of national strategies** (ST/ESA/245-UN, NYC, 1996). We

immediately recognized that these guidelines provided the missing elements in the U.S. approach to suicide prevention. A nine-month feasibility study followed. All who were interviewed agreed the model could be useful in the U.S. Since no organization offered to incorporate the guidelines, it was obvious that it would become the work of the grassroots—survivors of suicide. Thus, the **Guidelines** became the base or foundation of SPAN USA.

The SPAN USA PLAN, Figure 1 (page 5), is the visual model of the mission of SPAN USA: **Dedicated to the creation and implementation of proven effective national suicide prevention strategies.**

UNDERSTANDING THE PLAN. Read the intertwined circles in a clockwise pattern. Start at the bottom, or **GRASSROOTS.** Grassroots represents the family or each community where suicide occurs and where suicide prevention must likewise take place.

ADVOCACY is the effort of the people to bring to the attention of their elected officials the need for resources to prevent suicide. It is based on the first amendment to the U.S. Constitution: "The people have a right to peaceably assemble and petition the government for change."

AWARENESS makes the public conscious of the need for suicide prevention. Every May there is a SPAN USA gathering in Washington, D.C., to demonstrate the need for suicide prevention legislation and to hand-deliver educational materials and signed petitions to each member of Congress. **Lifekeeper** State "Faces of Suicide" Quilts that portray real people who have died by suicide are displayed.

ADVOCACY and **AWARENESS** work together to develop "political will." The combination focuses attention to the problem of suicide and moves legislators to action!

RECOGNITION has been made of advocacy efforts by the Congress (1) in unanimously passing S. Res. #84 and H. Res. #212 (105th Congress) that declare suicide a national problem; (2) in conducting a Senate Hearing on Suicide (106th Congress); and (3) in appropriating $3 million for crisis line evaluation and authorizing $75 million for suicide prevention (106th Congress).

A 1997 **STEERING** committee provided the direction for a Consensus Development Conference in the five high-suicide risk areas: the young; the elderly; consumers of mental health services; the chronically ill and a diverse population of Hispanic, Native American/Alaskan Indians; young black males and the gay, lesbian, transsexual, and bisexual people.

The SPAN USA **SUMMIT** meeting in Reno, Nevada, October 1998, presented the meta-analyses, peer-reviewed data of the best suicide prevention practices in the five high-risk areas (Tool #1a). Presented to Assistant Secretary of Health and Human Services/ Surgeon General David Satcher were the 81 recommendations for suicide prevention distilled by an expert panel from more than 700 concepts presented by the conference delegation.

With the guidance of Dr. Satcher, a federal-private partnership extracted fourteen final recommendations to begin a national Suicide Prevention **STRATEGY**. These were published as The Surgeon General's Call to Action to Prevent Suicide 1999 (Tool #1b).

The **STRATEGY** was further developed into Goals and Objectives of the National Suicide Prevention Strategy by Department of Health and Human Services experts (Tool #2). Yet to be produced are the Action Steps needed to implement this **STRATEGY**.

This booklet: ***SUICIDE PREVENTION: PREVENTION EFFECTIVENESS and EVALUATION*** (Tool #3), provides the basic tools needed to create effective suicide prevention plans. State Suicide Prevention Plans can adapt these evidence-based and best practice recommendations to meet the needs of the people in their states. This process allows needed resources to be brought back to the **GRASSROOTS**.

But all this effort cannot lay dormant at the grassroots. It is only effective if it is implemented, **EVALUATED, SHARED** with others, and **MODIFIED** to meet the changing needs of each community. Only then can we begin to realize the value of **Linking Research to Practice—the theme of the Reno Conference—AND OF SAVING LIVES LOST TO SUICIDE!**

Gerald (Jerry) and Elsie Weyrauch, August 3, 2001

Suicide Prevention:
Prevention Effectiveness and Evaluation

Introduction

This booklet explains important prevention and evaluation concepts in the context of suicide prevention. It is designed for prevention program managers and staff to facilitate a common understanding and language with prevention and evaluation specialists. Reading the booklet will not prepare someone to assess prevention effectiveness or to evaluate a program without consulting an authority on these topics. The booklet will, however, help readers understand the need for applying the principles of prevention effectiveness and for incorporating evaluation into program planning and implementation. Both prevention effectiveness and evaluation are the keys to preventing suicide. Persons working to prevent suicide need a basic understanding of the keys to know that they are do-able, available, worthwhile, and *essential* if suicide prevention is to move forward and lives are to be saved.

Section 1
Principles of Suicide Prevention Effectiveness

Suicide has been a leading cause of death in the United States for years, yet we understand little about the actual causes of suicide. We understand even less about how to prevent suicide. Many people are trying to develop and implement suicide prevention efforts. Application of basic principles of prevention effectiveness would benefit these efforts by helping them to maximize their impact. An understanding of the concepts and principles relevant to prevention effectiveness that are presented in this document should promote their application in practice.

Even though suicide is a leading cause of death in the United States, the Federal financial resources applied to preventing suicide are dismal at best. The U.S. Department of Health and Human Services currently spends an estimated 20 million dollars annually on this problem. Much of this funding goes to research and not programs. Other causes of death that take less of a toll than suicide receive billions of dollars in annual funding.

Collectively, States are probably spending more on suicide prevention programs than the Federal government. This itself is a testament to the fact that limited resources must be used in the best possible way. And this will require that all of us do the very best with what we have.

The enthusiasm, energy, and dedication that survivors of suicide bring to the prevention effort is one of our most significant resources. However, this grassroots group, even with the "fire in the belly" to prevent suicide, faces a tremendous lack of financial resources for suicide prevention. Thus, we must use our limited resources in the best possible ways. This means that planners of state and local community suicide prevention programs must understand and apply the basic principles of prevention effectiveness. To do less would be to waste the valuable and very limited resources we have.

Need for Information About Effective Prevention of Suicide

Prevention must be based on the most sound and best evidence available. Suicide prevention must begin with identifying prevention strategies, followed by research to

determine if these strategies work. Whenever we believe we have an effective strategy, we should explore the impact and cost of that strategy in a community setting, and then work to improve the strategy and its delivery.

Box 1
Principles of Suicide Prevention Effectiveness

- Prevention programs should be designed to enhance protective factors. They should also work toward reversing or reducing known risk factors. Risk for negative health outcomes can be reduced or eliminated for some or all of a population.
- Prevention programs should be long-term, with repeat interventions to reinforce the original prevention goals.
- Family-focused prevention efforts may have a greater impact than strategies that focus only on individuals.
- Community programs that include media campaigns and policy changes are more effective when individual and family interventions accompany them.
- Community programs need to strengthen norms that support help-seeking behavior in all settings, including family, work, school, and community.
- Prevention programming should be adapted to address the specific nature of the problem in the local community or population group.
- The higher the level of risk of the target population, the more intensive the prevention effort must be and the earlier it must begin.
- Prevention programs should be age-specific, developmentally appropriate, and culturally sensitive.
- Prevention programs should be implemented with no or minimal differences from how they were designed and tested.

Policymakers and program funders use information about effectiveness to set policy and funding priorities for suicide prevention. They want to implement suicide prevention strategies that work, and they generally consider the following issues when making decisions:

- Potential to reduce or avoid self-injury or death
- Social, legal, and ethical impact
- Economic impact
- Best methods to implement

Each of these considerations is crucial to a successful suicide prevention effort. It is best to consider these elements before starting a program. However, it is never too late to consider these elements within existing programs. Many suicide prevention programs lack evidence about effectiveness and could use the valuable information gleaned by evaluation to make immediate improvements in the program.

Prevention Targeting

In suicide prevention, "targeting" refers to a focus on modifying something that is causally related to self-injury within a specific population. The goal is to break the causal chain and prevent self-injury. There are two complementary ways to think about prevention targeting. One is to focus on the *level* of the prevention target. The other is to focus on the *injury stage*.

Target levels

One way to think about prevention is to focus on the level of the intervention target. A commonly used framework for describing intervention levels is "indicative," "selective," and "universal":

- *Indicated interventions* are highly targeted and frequently involve identification, treatment, and skill building among individuals and families. At this level the focus is on early detection and, frequently, intensive individual treatment relying on one-on-one, provider-to-patient interaction. Early-treatment programs are examples of indicated strategies. Indicated interventions tend to occur within the traditional health and mental health care delivery system and tend to be resource-intensive per person served.

- *Selective interventions* are targeted at high-risk groups with a focus on screening and group prevention activities. Peer support programs for students with a number of risk indicators would be an example. Selective interventions are less resource-intensive than indicated programs.

- *Universal interventions* are targeted at communities or larger aggregations and may include media or educational campaigns and other broad-population-based prevention strategies. Universal interventions may also be environmental prevention strategies that focus on physical changes that reduce risk. In suicide prevention, barriers on high places, modifications to automobile exhaust systems, and reducing access to lethal means, especially firearms, are examples of environmental strategies.

Targeting by stage

In suicide prevention, there are three stages of prevention—primary, secondary, and tertiary—which correspond to "before suicidal behavior occurs, as suicidal behavior occurs," and "after suicidal behavior occurs," respectively. Intervention strategies will vary, depending on the stage being targeted.

- *Primary prevention* refers to an effort that targets the causes of suicide-related behavior and injury before self-injury or suicidal behavior occurs. Such conditions as depression, impulsive behavior, or alcohol and drug abuse might be targets of primary prevention, usually through health and mental health services. Implementing programs that prevent alienation or isolation of youth, such as bullying prevention and improving access to health and mental health care for individuals, are other examples of primary prevention strategies. Finally, primary prevention strategies also include efforts to reduce access to lethal means or to address media coverage of suicide.

- *Secondary prevention* attempts to target intervention as behavior is occurring, with the goal of minimizing any self-injury that may occur. Early detection of suicidal ideation, or planning, and appropriate referral and treatment for suicide risk are examples of secondary prevention.

- *Tertiary prevention* targets intervention following self-injury or behavior to minimize the impact and reduce the likelihood of subsequent self-injury. Effective intervention in a suicidal crisis, therapeutic treatment following suicidal

behavior to prevent future attempts or to reduce the severity of an injury are examples of tertiary prevention. Referral for other supportive services following a suicide attempt is another example.

By combining the level of intervention and the stage of intervention we can describe and compare different strategies. The efficacy, effectiveness, and cost of each prevention strategy are important when making such comparisons. For example, comparing the effects of screening and treatment for depression (indicated, primary) with training gatekeepers (selective, secondary), or comparing a particular clinical treatment for depression (indicated, primary) with a strategy to reduce access to firearms among youth (universal, primary), can provide sound data to help us choose among and combine different strategies.

The key is to consider the effectiveness of various strategies for reducing suicide. With this information we can make decisions about the best use of scarce resources.

Assessing Prevention Effectiveness

Assessing prevention effectiveness is a scientific approach for making sure what we are doing or want to do will work.

Basic steps in assessing prevention effectiveness include the following:

- Identifying which strategies will be most likely to reduce injury and death from suicidal behavior
- Determining the potential effects of those strategies, including social, legal, ethical, and economic factors
- Determining optimal methods for implementing those strategies
- Assessing the effectiveness of a strategy periodically as it develops and is implemented

There are several aspects of prevention effectiveness that are important to understand. These include efficacy, effectiveness, safety, and economic analysis.

Efficacy: Does it work under ideal circumstances?

The first question that should be addressed about any suicide prevention technique is, "Does it work?" What is the evidence that justifies using the technique? How good is that information? *Efficacy* is defined as the effect obtained with a specific technique in expert hands under ideal circumstances. Determining efficacy of a prevention strategy requires a review of studies to examine the scientific evidence behind it and the potential magnitude of its impact.

Studies using experimental designs produce some of the most credible information about efficacy. In a standard experimental design, subjects are randomly placed into one of two similar groups. One group receives the program, while a control group does not. By comparing outcomes of the two groups, we can see the effects of the program.

Thus far, efficacy of most suicide prevention strategies has not been determined by using randomized trials. These types of studies are very costly and difficult to conduct. Smaller or methodologically less desirable studies often must be used to assess efficacy. At this point, most suicide prevention efforts currently in place assume efficacy, with little or no scientific evidence.

Effectiveness: Does it work in the real world?

Once we assume or confirm the efficacy of a suicide prevention strategy, we must ask, "How well does it work in the real world?" *Effectiveness* is the impact of the prevention activity in the real world. Effectiveness may be thought of as efficacy of a strategy as assessed in the hands of practitioners within real-world constraints. Under experimental conditions, real-world constraints are minimized. However, in real-world applications practical difficulties may keep the program from being effective. For example, a theoretically sound program may have difficulty getting people to participate. To adjust for real-world constraints, effectiveness studies are done in the setting in which the intervention will be conducted, often in community demonstration projects.

Safety: How safe will it be in the real world?

Do no harm is an ethical principle that should be at the forefront of concern when implementing any program. Some suicide prevention techniques are associated with potential hazards. Some hazards may result directly from prevention efforts. For

example, anti-depressant medications may cause side effects or drug interactions. Other hazards may result from prevention measures; for example, vulnerable youth may be distressed from exposure to certain suicide prevention education curricula.

Initial safety data emerge from efficacy studies. Additional information about safety is generated when this kind of evaluation is applied on a broader scale. Potential safety risks must be assessed before and during any prevention program. Safeguards should always be in place in any program to ensure that risks of harm are small or eliminated. When pilot-testing or developing strategies with input from vulnerable populations, it is imperative that their safety is considered and that appropriate safeguards are in place.

Economic studies: How cost-effective is it?

Economic studies allow us to compare the costs and the benefits of a strategy. *Cost-effectiveness* refers to the dollars spent for each unit of health improvement, for example, dollars per suicide prevented. *Cost-benefit* analyses consider how much society values the outcome or is willing to pay for the outcome. This requires placing a value on various states of being. For example, we would need to assign a dollar value to the cost of a suicide. Then, with information on how much it would cost to prevent a suicide, we could determine the relative value of investing in suicide prevention. Efforts to conduct cost analysis represent a potential way of measuring the effectiveness of the prevention strategy.

A first step in economic analysis is to determine the total program cost, including direct and indirect costs. Direct costs include personnel, equipment, and space. Indirect costs include costs of time to the recipient of the program, lost time from work, and travel. Direct benefits include costs saved from avoiding the outcome (e.g., health care costs). Indirect benefits include costs saved (e.g., earned wages and productivity).

With information about efficacy, effectiveness, safety, and cost, strategies can be compared and decisions made about how to best invest resources. However, for suicide prevention, we are a long way from being able to systematically make such comparisons. We are left with doing the best we can to prevent suicide with limited information about efficacy. However,

existing and new program efforts can and must make an effort to contribute to the information we need for more effective prevention of suicide.

Principles of Suicide Prevention Effectiveness

There is limited information about the efficacy and effectiveness of any suicide prevention strategy. Yet suicide prevention efforts are ongoing around the world. Void of information about efficacy, we can work to deliver programs effectively by reducing constraints to program delivery. A number of principles of effectiveness from drug prevention strategies have been adapted for implementing suicide prevention efforts (see Box 1).

Unfortunately, very few suicide interventions have been thoroughly evaluated for efficacy and safety. Because suicide prevention programs have been implemented before appropriate assessments are completed, gaps may exist in knowledge about the efficacy, effectiveness, safety, or economic impact of specific prevention strategies.

Section 2
Evaluating Suicide Prevention Programs

Lack of sound evaluation remains one of the most significant barriers to identification and implementation of effective intervention strategies. Evaluation has the potential to produce information on program efficacy and effectiveness and, at the very least, to provide information that will improve delivery of programs.

Purpose of Evaluation

The single greatest obstacle to effective prevention of suicide is the lack of evaluation research. Evaluation of prevention programs ensures the best use of limited resources. Questions from legislators, professionals, funders, scientists, and survivors about the use of resources can be addressed with the critical information produced by evaluation. For example, what contribution did the application of these resources make to suicide prevention? How does the contribution of one program measure up to the contribution made by other programs? By answering these questions, communities and States are better informed to make evidence-based choices when selecting and implementing programs.

The following pages should serve as a guide to program evaluation for persons working in the area of suicide prevention. The emphasis is on practical, ongoing evaluation strategies that involve all program stakeholders, not just evaluation experts.

Evaluation is an important part of the SPAN plan (see Figure 1). As professionals *evaluate* programs, they begin *sharing* what works best. Through evaluation and sharing, programs are *modified* and adapted for use in various settings.

Evaluation is easier than most people believe. A well-designed and well-run suicide prevention program produces most of the information needed to determine its effects. The key to success for effective evaluation is preparation. The ease of evaluating a program depends on the effort put into program design and operation. Tension often develops between spending resources on service delivery and on evaluating the program. However, programs that can demonstrate effectiveness and efficiency are more likely to obtain legislative, community, technical, and financial support.

Program evaluation is a way to help suicide prevention efforts be more effective. Evaluation is the process of determining how well programs work. Evaluation can identify benefits and problems of a program. Evaluation information can improve the delivery of effective programs. **Without evaluation of programs, we do not know if the program benefits or harms the people we are trying to help.**

Evaluation tells stakeholders if the program is achieving its goals and if the program needs to be modified. Additionally, evaluation can improve the morale of program personnel, as program staff see that their efforts are not wasted and develop and implement strategies for addressing needs identified by the evaluation.

Evaluations frequently produce unexpected, but useful, information, either about something that works or about something that needs improvement. Evaluation helps to communicate aspects of your program to other agencies or groups, especially if it is published in a scientific journal or a more informal medium. As described in Box 2, there are many benefits of program evaluation.

Box 2
Benefits of Program Evaluation

- Learning whether proposed program materials are suitable
- Learning whether program plans are feasible
- Providing an early warning system for problem identification
- Learning whether programs are producing the desired results
- Learning whether programs have any unexpected benefits or problems
- Enabling managers to improve service
- Monitoring progress toward the program's goals
- Producing data on which to base future programs
- Demonstrating the effectiveness of the program
- Identifing the most effective parts of the program for refinement
- Gathering valuable information that can be shared

Cost of Evaluation

The cost of evaluation varies as a result of a number of factors, including experience and education of consultants, the type of evaluation required, the population density or size, and the geographic location of your program. Generally, about 15 to 20 percent of available program funds should be budgeted for evaluation. The exception is for a program evaluation with an experimental design. In most experimental designs, the cost of evaluation will be extensive. The cost of evaluation should always be included in proposals for grant funds. Resources spent on evaluation should be viewed as an investment in future success, with dividends of reduction in suicidal behaviors.

Selecting an Evaluator

Program personnel need the help of an outside evaluator who is hired to focus on evaluation. The best and most appropriate evaluators are those with no personal interest in the results of an evaluation. In most cases, outside consultants are best. They will look at the program from a new perspective. There are a number of places to identify potential evaluation consultants; for example, universities, local foundations, prevention agencies, and nonprofit and for-profit companies frequently have evaluation professionals that can provide consultation.

It is important to consider the professional training and experience level of consultants. Not all evaluators are the same. Some work on community evaluation, while others specialize in other types of evaluation, such as policy impact. Some specialize in quantitative methods, others in qualitative methods. Some have experience with one stage of evaluation, others with another stage. It is important to find a consultant whose background and training best fit the program's evaluation goals.

Integrating Evaluation into the Program

Evaluation should be a routine part of program operations. Evaluation should involve program staff and stakeholders in gathering information for improving the program. Most routine evaluation efforts can guide small changes in programs. With appropriate resources, pilot-testing, and good record keeping, information to evaluate the effects of a program will develop naturally. When the purpose of the evaluation is to assist in making significant decisions, evaluation procedures need to be more extensive and formalized.

Criteria for assessing the program's success should be determined as part of planning before the program is implemented. The following aspects should be considered:

- A pilot-test of all the program's plans, procedures, activities, and materials
- Criteria for knowing if the program is working and being delivered as planned
- A system for gathering information

Components of Evaluation

Every evaluation plan must contain certain basic components. These include a clear objective, a description of the target population, and a description of what is to be evaluated.

Clear objectives

Writing a clear objective involves defining what is being evaluated. The objective will depend on the part of the program that is being evaluated. Without such a statement, evaluators are unfocused and do not know what to measure. For example, before the program begins, you will need to test any materials you plan to distribute to program participants. In such a case, your evaluation objective might read something like this:

> To learn whether the people in our target population can understand our new brochure about the risk factors associated with suicide

Your evaluation objective for a program component might read like this:

> To measure how many interventions were conducted or referrals made as a result of our gatekeeper training program

Description of the target population

In defining the target population, you should be as specific as possible. The target population will vary depending on the reason for the evaluation. A sample definition of a target population might read like this:

> All children in grades 10 through 12 who have been disciplined for truancy or alcohol or drug use and who attend public schools in the county

Description of what is to be evaluated

You will also need to clarify the type of information to be collected and how that information relates to your program's objectives. For example, if the goal of your program is to increase the number of referrals made to a mental health clinician by a person trained as a gatekeeper, you would need to collect information on the following:

> Number of referrals made and number of persons who seek services from the mental health clinician

Defining these components takes time and thought. Consulting an authority on program evaluation may be helpful. Once these basic components are in place, the evaluation effort can proceed.

Six Steps to Program Evaluation

There are six basic steps to program evaluation, all of which are related. The first steps provide a foundation for the later ones. The steps (Box 3) are: engage stakeholders, describe the program, focus the evaluation design, gather credible evidence, justify conclusions, and ensure use and share lessons learned.

Engage stakeholders

Most suicide prevention efforts involve partners, all of whom are stakeholders in the effort. Stakeholders must be involved in the evaluation; without their involvement, an evaluation might miss important parts of a program. Three groups of stakeholders are important:

- Those involved in program operations (e.g., sponsors, collaborators, coalition partners, funding officials, administrators, managers, and staff)
- Those served or affected by the program (e.g., clients, family members, neighborhood organizations, academic institutions, elected officials, advocacy groups, professional associations, skeptics, opponents, and staff of related or competing organizations)

Box 3
Summary of the Six Steps to Program Evaluation

1. **Engage stakeholders**—identifying the information they need will drive the evaluation
2. **Describe the program**
   a. Statement of need—why is the program needed?
   b. Expected effects—how does the program intend to address the need(s)?
   c. Activities—what does the program do?
   d. Resources—what does the program have that will enable its activities?
   e. Stage of development—how far along is the program in addressing the need(s)?
   f. Context—what is the environment of the program?
   g. Logic model—planned sequence and design of the program?
3. **Focus the evaluation design**
   a. Purpose—what are the objectives of the evaluation?
   b. Users—who's consuming the evaluation output?
   c. Uses—what are the users' information needs?
   d. Questions—what information will address users' needs?
   e. Methods—how will the information be collected, analyzed, and reported?
   f. Agreements—who's going to do what, and when?
4. **Gather credible evidence**
   a. Indicators—what information will address questions?
   b. Sources—where will the information come from?
   c. Quality—how good is the information?
   d. Quantity—how much information is needed?
   e. Logistics—what are the systems for collecting and managing information?
5. **Justify conclusions**
   a. Standards—what do we compare evaluation information to?
   b. Analysis and synthesis—how do we summarize and organize the information?
   c. Interpretation—how do we make sense of the information?
   d. Judgment—how do we compare findings to the standards?
   e. Recommendations—what should we do with our findings?
6. **Ensure use and share lessons learned**
   a. Design—think through the whole evaluation
   b. Preparation—plan for the evaluation and dissemination
   c. Feedback—communicate with users
   d. Follow-up—help users interpret findings and recommendations
   e. Dissemination—the most important part of a good evaluation

- Those who will be the primary consumers of information produced by the evaluation

Stakeholder involvement is different for each program evaluation. The objective of the evaluation will help define which stakeholders to involve. Without agreement from stakeholders, the evaluation may be of limited use.

Describe the program

The program description details the mission and objectives of the program. Descriptions should help evaluators understand the program goals and strategies. The description should discuss how the program works to effect change, describe the program's stage of development, and show how the program fits into the larger organization and community. Stakeholders should review and agree with the program description. Program descriptions will vary for each evaluation.

Aspects to include in a program description are a statement of need, expected effects, activities, resources, stage of development, context, and a logic model. A *statement of need* describes the problem that the program addresses. *Expected effects* are what the program must do to be successful. *Program activities* are what the program does to effect change. *Resources* include the time, talent, technology, equipment, information, money, and other assets available to conduct program activities. The program's *stage of development* reflects its maturity. The *context* should describe the setting within which the program operates. Some programs have used a *logic model* as a planning tool to outline resources and to describe the order of the program activities and associated outcomes. Logic models require a planned, sequenced thought process to design a program.

Focus the evaluation design

The evaluation must focus on issues of greatest concern to stakeholders. A funder may be concerned about efficient use of resources. The designer of an intervention may be concerned about fidelity of delivery. Most evaluations have limited time and resources. A good evaluation design is directly related to knowing its intended uses, that is, whether it is designed to demonstrate effectiveness to policymakers, acquire resources, or some other type of use. Significant items to consider in focusing

an evaluation are purpose, users, uses, questions, methods, and agreements.

Describing the *purpose* of the evaluation (i.e., its intent) will guide how it should be conducted. The stage of a program's development will define the purpose of the evaluation. For example, a program just starting may focus its evaluation on perceptions of the use of the materials being developed. A more developed program may attempt to conduct an outcome evaluation to see if the program is effective. The context of the program also affects the purpose. Evaluations of prevention programs have four general purposes. The first is to gain insight: We try to learn if doing something makes sense. A second purpose for program evaluation is to change practice. A third purpose is to assess the effects of the program. The fourth purpose is to effect change in the persons participating in the management and delivery of the program. Participating in an evaluation can provide insight for program staff that can be a catalyst for self-directed change.

Users are the persons who will receive evaluation findings, such as the program manager, funders, and other stakeholders. Intended users should participate in choosing the evaluation focus.

Uses are the ways in which information from the evaluation will be applied, as in deciding to continue funding or expand an effort or determining if program procedures should be modified for better delivery. Stakeholders should help identify, plan, and prioritize uses in advance of implementing the evaluation process. The program's stage of development and current context should also be considered when identifying uses.

Evaluation questions clarify what aspects of the program will be addressed. Creating evaluation questions identifies what the evaluation should answer. The questions will suggest how and what information should be gathered to enable adequate answers.

The *methods* for an evaluation refer to scientific research options. Methods for suicide prevention evaluations are generally developed in the social, behavioral, and health sciences and include experimental, quasi-experimental, and observational

designs. Choosing the best design depends on the circumstances. The method should enable the evaluation to address stakeholders' questions. *Experimental designs* randomly assign persons to intervention and non-intervention groups. By using random assignment, the two groups are usually very similar, with the exception that one receives the intervention and the other receives something else. When the intervention is complete, the differences we see in the two groups in the outcome of interest is likely caused by the difference made by the intervention. *Quasi-experimental methods* are similar to experimental, except that the groups are not similar because they are not randomly assigned. This type of method is used when randomization is not feasible. *Observational methods* tend to be descriptive and attempt to understand differences, similarities, and processes within a group. Monitoring a sample of telephone calls on a crisis telephone line and describing how well procedures are being followed is an example of an application of an observational method.

The choice of design has implications for what will be used as evidence. The design will determine how evidence will be gathered and what kind of claims can be made. The design also determines how data sources will be selected, what data collection instruments will be used, who will collect the data, and what data management systems will be needed.

Each evaluation method has its own limitations. Evaluations that mix methods are generally more effective. Methods might need to be revised or modified over the course of an evaluation: The intended use of an evaluation might shift, or changing conditions might require redesign of methods to keep the evaluation on track.

Agreements are explicit written statements that summarize the procedures and clarify roles and responsibilities among those who will execute the evaluation plan. Written agreements ensure understanding among stakeholders and evaluators in terms of their expectations and help to avoid numerous problems that can develop from misunderstandings.

To focus an evaluation design, stakeholders should review the evaluation questions and determine if they will address concerns of interest groups. A list of evaluation uses could be

circulated among stakeholders to determine which is most important. Intended users could also be interviewed to understand what they need or want to know and to focus the evaluation so it responds to those needs.

Gather credible evidence

The information collected from a suicide prevention evaluation should provide a useful understanding of the program that addresses stakeholders' needs. Stakeholders must find the information believable and relevant. Having credible evidence strengthens the recommendations. Credibility can be improved by using multiple procedures and by involving stakeholders in defining and gathering data. The following factors affect people's perceptions of the credibility of your evaluation evidence: indicators, sources, quality, quantity, and logistics.

- *Indicators* are aspects of the program that can be examined to address the questions of the evaluation. Examples of indicators that can be defined and tracked include the program's capacity to deliver services, the participation rate, levels of client satisfaction, the efficiency of resource use, and the amount of intervention exposure. Other measures of program effects, such as changes in participant behavior, community norms, policies or practices, health status, quality of life, and the settings or environment around the program, can also be tracked.

- *Sources* of evidence are persons, documents, or observations. More than one source might be used to gather evidence. Use of multiple sources provides different perspectives. In an evaluation where the question is, "Does a training of trainers result in a desired number of new gatekeepers being trained?", the indicator might be a "number of persons trained by each trainer over the course of a year."

- *Quality* refers to the correctness and integrity of the information. Quality data are representative of what they intend to measure and are informative for their intended use. Good indicators make it easier to collect quality data. Instrument design, data-collection procedures, training of data collectors, source selection, coding, data management, and routine error-checking all influence the quality of your

data collectors, source selection, coding, data management, and routine error-checking all influence the quality of your data. For example, a data collection effort where data collectors are trained on procedures and where there are checks to monitor adherence to procedures will produce higher-quality data than one with inconsistent or minimal training of data collectors.

- *Quantity* refers to the amount of evidence gathered. The amount of information needed should be estimated in advance. All evidence collected should have a clear and anticipated use, with only minimal burden placed on respondents. Information gathered from participants in an intervention should be kept to a minimum. Each indicator should be clearly justified in terms of addressing an evaluation question. Another example of quantity would involve determining how many persons must provide information to adequately address the evaluation question. The burden on persons to provide information should always approach the minimum needed.

- *Logistics* encompass the methods, timing, and physical infrastructure for gathering and handling evidence. Each technique selected for gathering evidence must be suited to the source(s), analysis plan, and strategy for communicating findings. Cultural issues should influence decisions about acceptable ways of asking questions and collecting information. Procedures for gathering evidence should be sensitive to cultural conditions in each setting and must ensure that the privacy and confidentiality of the information and sources are protected.

Justify conclusions

Evaluation conclusions must be drawn from the evidence gathered and then compared to the standards set by the stakeholders. Stakeholders must agree that conclusions are valid: otherwise, the evaluation results will be of limited use. Justifying conclusions on the basis of evidence requires the following: standards, analysis and synthesis, interpretation, judgment, and recommendations.

- *Standards* reflect what stakeholders think is important and are the basis for forming judgments concerning program performance. Using standards distinguishes evaluation from other approaches, in which priorities are set without reference to any sort of specific statement about what is important.

- *Analysis and synthesis* of an evaluation's findings might detect patterns in evidence. Analysis involves isolating important findings, while synthesis involves combining sources of information to reach a larger understanding.

- *Interpretation* is the effort of figuring out what the findings mean. It increases understanding of the evidence gathered in an evaluation. Merely uncovering facts regarding a program's performance is not sufficient to draw evaluative conclusions; one must interpret the results based on criteria set before the evaluation begins.

- *Judgments* are statements concerning the merit, worth, or significance of the program. They are formed by comparing the findings of and interpretations regarding the program against one or more selected standards. Because multiple standards can be applied to a given program, stakeholders might reach different or even conflicting judgments. For example, a program that increases referrals to a mental health clinician by 10 percent from the previous year might be judged positively by program managers who are using the standard of improved performance over time. However, community members might feel that, despite improvements, a minimum threshold of access to services has not been reached.

- *Recommendations* are actions for consideration resulting from the evaluation. Recommendations that lack sufficient evidence or that deviate from stakeholders' values can undermine an evaluation's credibility. Sharing early drafts of recommendations and asking for feedback from multiple stakeholders during the process can increase the likelihood that the recommendations will be accepted. Additionally,

when possible, presenting options instead of directive advice will make recommendations more acceptable.

Adequate justification of the conclusions and recommendations of an evaluation is an essential part of creating information that is useful. However, having useful information does not ensure that it is actually used.

Ensure use and share lessons learned

Unfortunately, despite the best intentions of evaluators and practitioners, lessons learned during an evaluation are not always used. For findings to be used effectively, they must be disseminated appropriately. This requires strategic thinking and should begin in the earliest stages of planning an evaluation and engaging your stakeholders. The goal of dissemination is to achieve full disclosure and impartial reporting. Additional uses for evaluation flow from the process of conducting it. Persons who participate in an evaluation can experience changes in thinking and behavior. Evaluation sometimes increases staff understanding of program goals.

Five critical elements for ensuring that an evaluation is used are as follows: design, preparation, feedback, follow-up, and dissemination.

- *Design* refers to the construction of evaluation questions, methods, and overall processes. The design should be organized to achieve intended uses by stakeholders.
- *Preparation* refers to sound evaluation planning and following the steps described in this booklet to ensure effective use of evaluation findings.
- *Feedback* is the communication that occurs among all parties to the evaluation. Giving and receiving feedback creates an atmosphere of trust among stakeholders.
- *Follow-up* refers to the technical and emotional support that users need both during the evaluation and after they receive evaluation findings.

- *Dissemination* is the process of communicating to others the lessons learned. The reporting strategy should be discussed with intended users and other stakeholders.

Sharing information learned from implementing and evaluating suicide prevention programs is one of the most important responsibilities we have in our effort to prevent suicide. Without sharing information, we are isolated in our efforts, and adoption of innovative and promising practices is limited. Effective diffusion of prevention information can have revolutionary effects on the practice of prevention efforts.

The most important aspect of applying lessons learned from evaluation is ensuring that the program is improved as a result. Recommendations should be specific regarding suggested shifts in program emphasis or in specific procedures. A specific written plan should be developed for implementing accepted recommendations. Subsequent evaluation cycles should examine the implementation of the changes and their impact.

Where Do We Go from Here?

Using prevention effectiveness and evaluation in suicide prevention programs is critical—but also challenging, in terms of time, effort, and cost. Getting started in using prevention effectiveness and evaluation through partnerships with others who are committed and experienced in these issues are your keys for future success. Begin in a small way and grow in your skills and actions. Build on your last effort and share your lessons with others so that the entire field can learn. **The potential success of our combined effort to prevent suicide begins with you, in your state and your community!**

Prevention Effectiveness and Evaluation Resources on the Internet

American Evaluation Association
http://www.eval.org/

CDC Evaluation Working Group
http://www.cdc.gov/eval

W.K. Kellogg Foundation Evaluation Handbook
http://www.wkkf.org/Documents/WKKF/EvaluationHandbook/default.asp

National Mental Health Association—Effective Prevention Programs
http://www.nmha.org/children/prevent/effective.cfm

Prevention First—Online Course on Basics of Prevention
http://www.onlinesyllabus.com/prevention/

Prevention Science & Methodology Group
http://www.biostat.coph.usf.edu/research/psmg/

Primer on Evaluation from the U.S. Department of Justice
http://www.bja.evaluationwebsite.org

Project STAR: Corporation for National Service
http://www.projectstar.org/star/AmeriCorps/ea_home.htm

Research-Based Prevention: A Pyramid for Effectiveness
http://www.cprd.uiuc.edu/levels.html

Taking Stock: A Practical Guide to Evaluating Your Own Programs
http://www.horizon-research.com/public.htm

United Way of America—Outcome Measurement Resources
http://national.unitedway.org/outcomes/publctns.htm

Youth Suicide Prevention Programs: A Resource Guide
http://wonder.cdc.gov/wonder/prevguid/p0000024/p0000024.asp

Parts of this publication were adapted from the following:

CDC. "A framework for assessing the effectiveness of disease and injury prevention." MMWR 1992; 41 (No. RR-3).

CDC. "Framework for Program Evaluation in Public Health." MMWR September 17, 1999; 48 (No. RR-11); 1–40.

Sloboda, Z. and David, S. "Preventing Drug Abuse Among Children and Adolescents." National Institute on Drug Abuse, 2001.

Thompson, N.J. and McClintock, H.O. "Demonstrating Your Program's Worth: A Primer on Evaluation for Programs to Prevent Unintentional Injury." Atlanta: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, 1998.

This document was produced by Lloyd Potter in consultation with Jerry and Elsie Weyrauch. Cheryl DiCara, Ellen Connorton, John Kalafat, and Susan Eastgard provided many sound comments and suggestions that have been incorporated into this document, though these individuals bear no responsibility for any deficiencies. Stephanie Malloy provided copy editing on an earlier version of this document and many useful suggestions.