RAND

NATIONAL DEFENSE
RESEARCH INSTITUTE

CHILDREN AND FAMILIES
EDUCATION AND THE ARTS
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INFRASTRUCTURE AND TRANSPORTATION
INTERNATIONAL AFFAIRS
LAW AND BUSINESS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
TERRORISM AND HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: Jump to Page 1

Support RAND

Browse Reports & Bookstore

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore the RAND National Defense Research Institute

View document details

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see RAND Permissions.

This report is part of the RAND Corporation research report series. RAND reports present research findings and objective analysis that address the challenges facing the public and private sectors. All RAND reports undergo rigorous peer review to ensure high standards for research quality and objectivity.

RAND NATIONAL DEFENSE RESEARCH INSTITUTE

Development and Pilot Test of the RAND Suicide Prevention Program Evaluation Toolkit

Joie D. Acosta, Rajeev Ramchand, Amariah Becker, Alexandria Felton

Prepared for the Office of the Secretary of Defense
Approved for public release; distribution unlimited

The research described in this report was prepared for the Office of the Secretary of Defense (OSD). The research was conducted within the RAND National Defense Research Institute, a federally funded research and development center sponsored by OSD, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community under Contract W74V8H-06-C-0002.

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2013 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (http://www.rand.org/publications/permissions.html).

Published 2013 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Preface

Evaluations are critical for assessing the impact of U.S. Department of Defense (DoD) investments in suicide prevention and can be used as the basis for decisions about whether to sustain or scale up existing efforts. The Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury asked the RAND Corporation to draw from the scientific literature and create a toolkit to guide future evaluations of DoD-sponsored suicide prevention programs (SPPs). The overall goal of the toolkit is to help those responsible for SPPs determine whether their programs produce beneficial effects and, ultimately, to guide the responsible allocation of scarce resources.

This report summarizes the three complementary methods used to develop the RAND Suicide Prevention Program Evaluation Toolkit; it is meant to serve as a companion to the toolkit itself and to provide additional background for those who are interested in learning about the toolkit's development. To develop the toolkit, RAND researchers examined the peer-reviewed literature and other evaluation toolkits and elicited feedback from program staff responsible for implementing DoD-sponsored SPPs. The toolkit is available at http://www.rand.org/pubs/tools/TL111.html.

The contents of this report will be of particular interest to DoD SPP managers and suicide prevention coordinators and should also be useful to policymakers in other sectors who sponsor or manage SPPs more generally.

This research was sponsored by the Office of the Assistant Secretary of Defense for Health Affairs and the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury and conducted within the Forces and Resources Policy Center of the National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community.

For more information on the Forces and Resources Policy Center, see http://www.rand.org/nsrd/ndri/centers/frp.html or contact the director (contact information is on the web page).

Contents

Preface ... iii
Figures and Tables ... vii
Summary ... ix
Acknowledgments ... xiii
Abbreviations ... xv

CHAPTER ONE
Introduction ... 1
Need for a Suicide Prevention Program Evaluation Toolkit ... 2
Purpose of This Report ... 3
Organization of This Report ... 4

CHAPTER TWO
Development of the Suicide Prevention Program Evaluation Toolkit ... 5
Literature Review ... 5
  Identifying Articles for Review ... 5
  Title and Abstract Review ... 6
  Full-Text Review ... 6
  Abstracting Consistent Information from Each Source ... 9
Review of Other Evaluation Toolkits ... 12
  Identifying the Key Components of Evaluation Toolkits ... 12
Outlining the Chapters of the Toolkit ... 14

CHAPTER THREE
Review and Feedback on the Toolkit by DoD Suicide Prevention Program Staff ... 17
Identifying Program Staff to Review the Toolkit ... 17
Eliciting Feedback from Program Staff ... 17
Program Staff Feedback ... 18
Toolkit Revisions Based on Program Staff Feedback ... 20

CHAPTER FOUR
Recommendations for Disseminating the Toolkit ... 23
Recommendations ... 23
  Continue to Refine the Toolkit ... 23
  Consider Converting the Toolkit to an Interactive Online Format ... 23

Continue Partnering with the SPARRC and Defense Suicide Prevention Office to Disseminate the Toolkit........ 23
Conclusion........ 24

APPENDIX
Feedback Form Used in the Pilot Test of the RAND Evaluation Toolkit for Suicide Prevention Programs........ 25

References........ 35

Figures and Tables

Figures

1.1. Suicide Rates, by Service, 2001–2010 1
2.1. Flowchart for Literature Search 9
2.2. Getting To Outcomes Ten Steps to High-Quality Prevention 13

Tables

1.1. Brief Summary of DoD Suicide Prevention Activities 2
2.1. Detailed Literature Search Strategies Used to Identify Sources 7
2.2. Inclusion and Exclusion Criteria 8
2.3. Evaluation Data Abstraction Form 10
2.4. Measure Data Abstraction Form 11
2.5. Review of Tools Offered in Existing Toolkits 14
2.6. Summary of Sources Used to Inform the Toolkit's Development, by Chapter 15
3.1. Toolkit Chapters Reviewed by Program Staff 18
3.2. Extent to Which Toolkit Chapters Met Their Objectives 19
3.3. Extent to Which Chapter Sections Were Clear and Tools Were Easy to Use 20

Summary

In response to increasing suicide rates among military personnel, the U.S. Department of Defense (DoD) has implemented many policy changes and invested in a number of programs to prevent suicide within its ranks. The DoD Task Force on the Prevention of Suicide by Members of the Armed Forces (2010) and a subsequent RAND report (Ramchand et al., 2011) both recommended that DoD evaluate existing programs and ensure that new programs include an evaluation component when they are implemented. Evaluations are critical for assessing the impact of DoD investments in suicide prevention and can be used as the basis for decisions about whether to sustain, scale up, or discontinue existing efforts. The Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury (DCoE) asked RAND to draw from the scientific literature and create a toolkit to guide future evaluations of DoD-sponsored suicide prevention programs (SPPs). The overall goal of the toolkit is to help those responsible for SPPs determine whether their programs produce beneficial effects and, ultimately, to guide the responsible allocation of scarce resources.

Purpose of This Report

This report summarizes the methods used to develop the RAND Suicide Prevention Program Evaluation Toolkit; it is meant to serve as a companion to the toolkit itself and to provide additional background for those who are interested in learning about the toolkit's development.

Methods Used to Develop the Toolkit

We used three complementary methods to develop the toolkit: a peer-reviewed literature review, a review of other evaluation toolkits, and reviews and feedback by program staff responsible for implementing DoD-sponsored SPPs.

Literature Review

We conducted an extensive examination of peer-reviewed suicide prevention evaluation studies and clinical trials of suicide prevention activities. The literature review consisted of two steps. First, we identified relevant sources through three phases of web-based searches of peer-reviewed literature in content-relevant databases. These sources underwent successive rounds of screening, including a title and abstract review, followed by a full-text review, to exclude irrel-

evant and unsuitable articles. We focused on sources that described the evaluation of suicide prevention or reduction programs.

Second, we abstracted information from each source selected for inclusion (n = 166). We divided these sources into two tiers. We then coded the first-tier articles, which included evaluation studies of SPPs, and abstracted in a consistent way the pertinent data on both the evaluation details and the outcome measures. Evaluation data included descriptions of suicide prevention/reduction programs, details about evaluation design, and synopses of study findings. Measure data described how study outcomes were assessed and included details about measure administration, scoring, and reliability. Each piece of abstracted information represents a characteristic or quality of a program or measure that was useful to consider when constructing the toolkit. The abstracted data on evaluation can be found in Appendix A of the toolkit, and the abstracted data on measures was used to populate Table 4.1 ("Sample Process Measures") and Table 4.2 ("Sample Outcome Measures") in the toolkit.

Second-tier articles—those that were relevant to the toolkit's development but inappropriate for data abstraction (e.g., articles that described an evaluation methodology but that were not actual evaluations of SPPs)—were subsequently reviewed by the first author, who used information from these articles to inform the appropriate sections of the toolkit.

Review of Other Evaluation Toolkits

We used other evaluation toolkits to identify key components that should be included in our toolkit and to develop the format of the toolkit. We first reviewed existing evaluation toolkits to develop an outline and determine the types of tools that they offered. Tools identified included sample measures, checklists with yes/no questions, worksheets with open-ended questions to help guide users through the toolkit (i.e., through planning an evaluation, analyzing data, and using evaluation data to improve the SPP), and designs of prior evaluation studies. We relied, in particular, on the Getting To Outcomes® (GTO) approach because GTO is currently the only evidence-based model and intervention that has been proven to increase programs' ability to conduct self-evaluations (Acosta and Chinman, 2011).

Review and Feedback by Program Staff

We developed an initial draft of the toolkit and shared it with 12 program staff responsible for implementing SPPs in the Air Force, Marine Corps, and National Guard. Program staff were asked to spend five to six hours reading through the toolkit and completing the worksheets, checklists, and templates. We asked them to compile their feedback using a standardized feedback form (see the appendix to this report). The feedback form asked the staff member to indicate the extent to which each chapter of the toolkit met its objectives, whether there were any sections that were not clear or otherwise difficult to understand, and whether he or she felt uncomfortable using the tools provided. Program staff emailed their completed feedback forms to the RAND team and participated in a follow-up conference call to discuss their feedback.

Revisions to Toolkit Based on Program Staff Feedback

Based on program staff feedback, we made several improvements to the toolkit. For example, we added process measures in the body of the toolkit, identified the potential cost of outcome measures, and provided guidance on how to get more information about potential evaluation measures. Although several program staff suggested that the toolkit be converted to a more

interactive online product, such a change was outside the scope of the current project. Therefore, no revisions of this nature were made to the toolkit.

Recommendations for Toolkit Dissemination

Based on program staff feedback, the RAND team made three recommendations for DCoE to consider as it develops plans for disseminating the toolkit.

1. *Continue to refine the toolkit.* Continued monitoring will help ensure that the content remains up-to-date and relevant to users.
2. *Consider converting the toolkit to an interactive online format.* This format would allow for an autofill feature to prepopulate information, making it unnecessary for program staff to transfer information by hand.
3. *Continue partnering with the Suicide Prevention and Risk Reduction Committee and the Defense Suicide Prevention Office to disseminate the toolkit.* Dissemination strategies could include webinars to introduce the toolkit, distributing copies of the toolkit at conferences that suicide prevention coordinators may attend, and distributing the toolkit to installation-level contacts via email.

Conclusion

The RAND Suicide Prevention Program Evaluation Toolkit was developed to build the knowledge and skills of individuals responsible for implementing DoD-sponsored SPPs as they self-evaluate those programs. Disseminating the toolkit across DoD will help ensure that ongoing and future SPPs have guidance in determining whether their programs are having beneficial effects and will ultimately help guide the responsible allocation of scarce resources.

Acknowledgments

We gratefully acknowledge the assistance of the U.S. Department of Defense Suicide Prevention and Risk Reduction Committee, whose members provided support during the pilot testing of the toolkit: Thomas Burke, Christopher Dorr, Colonel John Forbes, Robert Gardiner, Jackie Garrick, Malcolm Hawkins, Joan Hunter, Jan Kemp, Leonard Litton, Lieutenant Commander Andrew Martin, Major Michael McCarthy, Walter Morales, Rebecca Porter, and Mark Reger. We also thank Keeley Judd for the administrative support she provided in preparing this document. In addition, we thank our project monitor at the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury, Captain Janet Hawkins, for her support of our work.

We benefited from the valuable insights we received from Sarah Hunter, Laurie Martin, and Commander Aaron Werbel. Their constructive critiques were addressed as part of RAND's rigorous quality assurance process and improved the quality of this report. Finally, we thank the staff from the 12 suicide prevention programs, convened as part of the pilot test, for sharing their time and feedback: Patti Busser, Anne Doty, Lieutenant Alicia Dudley, Joan Fischer, Major Kim Floyd, Andrea Gonzalez, Major Elizabeth Najera, Shanon Neal, Major Richard Soto, Lieutenant Commander Kenneth Simmet, Adam Walsh, Lieutenant Sara Wright, Mark Ybaben, and Captain Emily Young. Finally, we offer posthumous thanks to Lynne Oetjen-Gerdes at the Office of the Armed Forces Medical Examiner and Karen McGruder at the RAND Library, both of whom provided important insight and support for this project.

Abbreviations

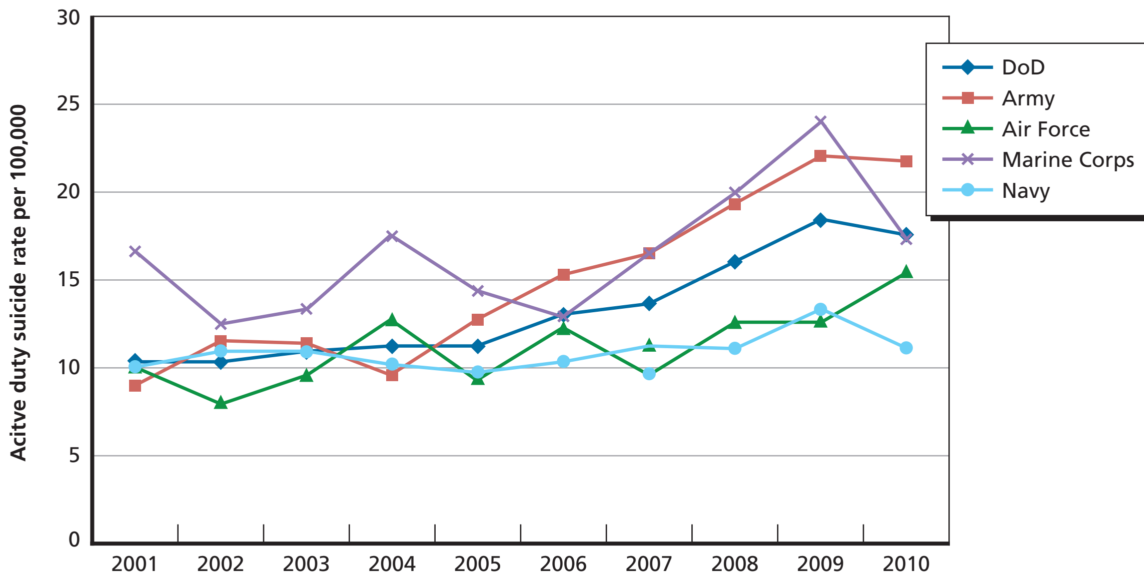
CQI	continuous quality improvement
DCoE	Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury
DoD	U.S. Department of Defense
GTO	Getting To Outcomes®
NREPP	National Registry of Evidence-Based Programs and Practices
SPARRC	Suicide Prevention and Risk Reduction Committee
SPP	suicide prevention program

CHAPTER ONE

Introduction

Stress on U.S. forces generated by multiple deployments to war zones has increased as a result of recent and ongoing combat in Afghanistan and Iraq. Suicide rates have also increased across all services, with the Army and Marine Corps showing the greatest rise since 2001 (Ramchand et al., 2011). In 2008, the suicide rate across the U.S. Department of Defense (DoD) was higher than it was between 2001 and 2005 (see Figure 1.1). Data from the 2008 DoD Survey of Health Related Behaviors Among Active Duty Military Personnel indicate that almost 6 percent of active-duty military personnel have attempted suicide, with 3 percent doing so since joining the military (Bray et al., 2009). A 2009 study found that close to 12 percent of active-duty military personnel reported seriously considering suicide in the past, with 3.3 percent doing so since joining the military (Bray et al., 2009).

Figure 1.1
Suicide Rates, by Service, 2001–2010



SOURCES: 2001–2008 data from the DoD Mortality Surveillance Division and the Office of the Armed Forces Medical Examiner (as of April 1, 2010); 2009 and 2010 data from Kinn et al., 2011.
RAND *RR283-1.1*

Need for a Suicide Prevention Program Evaluation Toolkit

In response to these suicide rates, DoD has been actively engaged in preventing suicides among service members. In accordance with Section 733 of the National Defense Authorization Act for Fiscal Year 2009 (Pub. L. 110-417), the DoD Task Force on the Prevention of Suicide by Members of the Armed Forces was convened "to examine matters relating to prevention of suicide by members of the Armed Forces." The subsequent task force report offered 76 recommendations critical to a successful DoD suicide prevention strategy, including the recommendation to "Ensure that all initiatives and programs have a program evaluation component" (section 7.4.3). The report indicated that many DoD suicide prevention efforts were not being evaluated and emphasized the need for program evaluation to improve knowledge about the effectiveness of any individual initiative and to encourage evidence-based investments of DoD resources, effort, and time.

A subsequent RAND report titled *The War Within: Preventing Suicide in the U.S. Military* (Ramchand et al., 2011) cataloged the suicide prevention activities across DoD and made recommendations to ensure that these activities reflected best practices. The study assessed how each of the services was performing across six domains that RAND identified as indicative of a comprehensive suicide prevention program (SPP) (see Table 1.1).

Table 1.1
Brief Summary of DoD Suicide Prevention Activities

Domains Indicative of a Comprehensive SPP	Suicide Prevention Activities
Raise awareness and promote self-care	All services used media campaigns, training and educational courses, and messages from key leaders to raise awareness of the signs and symptoms associated with suicide. Fewer messages were focused on self-care and were directed primarily at deploying personnel or those returning from deployment.
Identify those at risk	The Army, Navy, and Marine Corps trained gatekeepers, such as peers, chaplains, and front-desk attendants, at installation gyms to identify individuals at increased risk for suicide and to actively refer those in distress for follow-up care. The Air Force monitored the aftermath of high-risk events (e.g., airmen under investigation) and monitored service members after deployment, the latter a strategy also employed by the Army and Navy.
Facilitate access to quality care	To facilitate access to care, services located behavioral health care facilities in nontraditional settings, including primary care (Army and Air Force) settings and within units in theater (Marine Corps and Navy). The Real Warriors media campaign seeks to promote help-seeking across DoD.
Provide quality care to those in need	Behavioral health care professionals in the Air Force and Marine Corps were trained in suicide risk assessment and management.
Restrict access to lethal means	No known specific policies were in place to reduce access to lethal means.
Respond appropriately to suicides and suicide attempts	The services had a team or other personnel whom leaders could call to assist them after a suicide or traumatic event.

SOURCE: Ramchand et al., 2011.

More detail about these activities can be found in the RAND report (Ramchand et al., 2011). However, the study made 14 recommendations for DoD to consider in its ongoing suicide prevention efforts. Similar to the DoD task force, one of the RAND report's recommendations was to evaluate existing programs and ensure that new programs include an evaluation component when they are implemented. Another recommendation was to evaluate gatekeeper training, in particular, because this type of training is a prevention technique for which there is little empirical evidence of effectiveness (Ramchand et al., 2011).

Evaluations are critical for assessing the impact of DoD investments in suicide prevention and can be used as the basis for decisions about whether to sustain or scale up existing efforts. For SPPs, such evaluations are challenging, in part because suicide is rare (i.e., suicide occurs at a rate of less than 25 per 100,000 active-duty service members), and finding statistically significant effects in suicide outcomes is difficult. In addition, many SPPs have multiple components, making it difficult to discern which components or characteristics are responsible for a given observed effect. As a result, the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury (DCoE) asked RAND to draw from the scientific literature to create a toolkit to guide future evaluations of DoD-sponsored SPPs.

Purpose of This Report

This report summarizes the methods used to develop the RAND Suicide Prevention Program Evaluation Toolkit and is meant to serve as a companion to the toolkit (available at http://www.rand.org/pubs/tools/TL111.html) and to provide background on the development of the toolkit. The overall goal of the toolkit is to help those responsible for SPPs determine whether their programs produce beneficial effects and to ultimately guide the responsible allocation of scarce resources. To accomplish this goal, the toolkit has four specific aims:

1. to inform users about the most current evaluation research from other similar SPPs
2. to help users design an evaluation that is applicable to each type of program and available resources and expertise
3. to support users' selection of measures for new evaluations and to augment or enhance ongoing evaluations
4. to offer users basic guidance on how to analyze evaluation data and then use these data to improve the effectiveness of SPPs.

To ensure that the toolkit accomplished the above aims, we used three methods to develop it. First, we conducted an extensive review of peer-reviewed suicide prevention evaluation studies and clinical trials of suicide prevention activities. Second, we reviewed other evaluation toolkits to identify key components that should be included in the toolkit and to devise an outline for the format of the toolkit. Third, we developed an initial draft of the toolkit and then shared it with 12 program staff who were responsible for implementing SPPs in the Air Force, Marine Corps, and National Guard. These individuals were instructed to use the toolkit and provide feedback on its usability. We then revised the toolkit based on this feedback.

Organization of This Report

Chapter Two provides more detail about how the initial draft of the toolkit was developed. Chapter Three summarizes how we elicited feedback on the initial draft of the toolkit from staff responsible for implementing 12 SPPs and describes the revisions made to the toolkit as a result. Chapter Four provides recommendations to DoD about how to disseminate the toolkit.

Development of the Suicide Prevention Program Evaluation Toolkit

To develop an initial draft of the toolkit, we conducted a systematic literature review of prior studies of SPPs and reviewed other evaluation toolkits. We used the review of other evaluation toolkits to develop an overall framework for our toolkit and to determine how best to integrate the literature review findings. In this chapter, we describe each step in detail.

Literature Review

First, we identified relevant sources through web-based searches of peer-reviewed literature in content-relevant databases. These sources underwent successive rounds of screening, including a title and abstract review, followed by a full-text review, to exclude irrelevant and unsuitable articles. We focused on sources that described the evaluation of suicide prevention or reduction programs.

Next, we abstracted information from each source selected for inclusion. We divided these sources into two tiers. We then coded first-tier articles, which included evaluation studies of SPPs, and abstracted in a consistent way the pertinent data on both the evaluation details and outcome measures. Second-tier articles—those that were relevant to the toolkit but were inappropriate for data abstraction (e.g., articles that described an evaluation methodology but that were not actual evaluation of SPPs)—were subsequently reviewed by the first author, who used information from these articles to inform the appropriate sections of the toolkit.

Identifying Articles for Review

To identify relevant sources, we conducted database searches in two phases. Phase 1 consisted of searching for evaluation studies used to verify SPPs as evidence-based (n = 44 sources). Phase 2 consisted of a more comprehensive literature search on evaluation and clinical trials of SPPs in databases of peer-reviewed literature.

Phase 1: Evaluation Studies of Evidence-Based SPPs

We located these phase 1 evaluation studies in the Substance Abuse and Mental Health Services Administration's National Registry of Evidence-Based Programs and Practices (NREPP) for interventions with "suicide" as a listed outcome keyword. Due to the small size and specific focus of NREPP, the only search criterion for phase 1 was that suicide prevention was an intervention outcome. In NREPP, we searched each intervention for "studies" listed in the "quality of research" section. We also searched the Suicide Prevention Resource Center's Best Practices Registry to identify evaluation studies used to verify Section I (evidence-based) programs. Section I programs included interventions that have undergone rigorous evalua-

tion and demonstrated positive outcomes. No searching was necessary in the Suicide Prevention Resource Center's Best Practices Registry because all programs focused on suicide prevention as an intervention outcome.

Phase 2: Database Searches

Phase 2 consisted of a more comprehensive literature search on SPP evaluation in five databases in substantive areas pertaining to health (psychology and medicine), defense, and the social sciences broadly: PsychINFO (psychology), PubMed (medicine), Defense Technical Information Center (defense), New York Academy of Medicine Grey Literature Report collection (medicine), and Social Science Abstracts (social sciences). We also searched specifically for SPP clinical trials in three of the aforementioned databases: PsychINFO, Defense Technical Information Center, and Social Science Abstracts. We conducted multiple searches in each database to enhance the breadth of the results. We restricted our searches to articles written in English and published in peer-reviewed journals. Search strategies varied for each phase of the process and depended on the constraints of each electronic database. In general, searches contained keywords combining "suicide," "program," and "evaluation," or "suicide" and "clinical trial." Our search was conducted in January–February 2012, and keywords were chosen following consultation with the study team. We did not limit our search to a specific range of years, but searched all years available in each database. Details of the search strategies can be found in Table 2.1. We found additional articles by reviewing the references in articles identified for inclusion in the review.

Title and Abstract Review

To help ensure that all sources identified in the literature searches in phase 2 were relevant, we reviewed the article titles to remove articles that were clearly irrelevant to the current project. For the remaining set, we reviewed abstracts and categorized the articles into three groups: articles that were clearly irrelevant (e.g., epidemiological studies of suicide), articles whose relevance was questionable, and articles that should move forward to full-text review. Questionable articles were discussed among the team, and a decision was made about whether to include or exclude them. Articles selected for full-text review were further categorized as either *first-tier* (empirical evaluations of suicide prevention or reduction programs) or *second-tier* (articles that did not include such an evaluation but were still relevant to the toolkit, such as discussions of evaluation methodology). All articles identified during phase 1 were included in the first-tier. An electronic record of inclusion and exclusion criteria was updated as decisions were made (see Table 2.2).

Full-Text Review

Articles identified for full-text review were also carefully examined for information that was relevant to the inclusion and exclusion criteria. During the full-text review, articles that met the inclusion criteria were then coded according to the process outlined in the following sections.

Articles Identified During the Literature Search

The phase 1 and 2 database searches yielded 484 unique sources. The title and abstract review identified 271 articles for exclusion. We obtained full-text versions of the remaining articles. Another 22 articles were excluded during the full-text review. The primary reason for exclusion during the full-text review was an article's failure to provide an actual evaluation of a

Table 2.1
Detailed Literature Search Strategies Used to Identify Sources

Database	Limits	Search Concepts	Results
SPP Evaluation			
PubMed	Free text search	Concept 1: suicide prevention program* AND Concept 2: evaluat*	433 articles After duplicates and foreign-language articles were removed, 151 articles were retained.
		Concept 1: (“suicide”[Mesh]) AND “prevention and control” [Subheading] AND Concept 2: “program evaluation”[Mesh] AND Concept 3: (Search using indexing)	
	Free text search	Concept 1: (suicide prevent*[Title] OR suicide awareness[Title] OR suicide intervent*[Title]) AND Concept 2: program*	
PsychINFO	Peer-reviewed	Concept 1: “suicide prevention” OR “suicide awareness” OR “suicide intervention” AND Concept 2: program* AND Concept 3: SU evaluat*	221 articles After duplicates and foreign-language articles were removed, 37 articles were retained.
	Peer-reviewed	Concept 1: SU (“suicide prevention” OR “suicide awareness” OR “suicide intervention”) AND Concept 2: program* AND Concept 3: SU evaluat*	
	Peer-reviewed	Concept 1: “suicide prevention” OR “suicide awareness” OR “suicide intervention” AND Concept 2: program* AND Concept 3: SU (assess* OR review OR evaluat*)	
	Peer-reviewed	Concept 1: SU (“suicide prevention”) AND Concept 2: program* AND Concept 3: SU evaluat*	
Defense Technical Information Center	Technical reports	Concept 1: ti:(suicide and [prevent* or intervention or awareness]) AND Concept 2: program* AND Concept 3: evaluat* or review OR effica*	25 articles 3 articles were retained.
New York Academy of Medicine Grey Literature Collection	Gray literature	Subject: suicide prevention	18 articles 1 article was retained.
Social Science Abstracts	Peer-reviewed	Concept 1: SU suicide prevention AND Concept 2: program* AND Concept 3: evaluat*	48 articles After duplicates and foreign-language articles were removed, 3 articles were retained.

Table 2.1—Continued

Database	Limits	Search Concepts	Results
SPP Clinical Trials			
PsychINFO	Peer-reviewed	Concept 1: DE “suicide” OR DE “attempted suicide” OR DE “Suicidal Ideation” AND Concept 2: DE “clinical trials” OR DE “treatment outcomes”	508 articles After duplicates, foreign-language articles, and off-target reports were removed, 216 articles were retained.
	Peer-reviewed; free text search	Concept 1: TI (suicid*) AND Concept 2: TX (prevent*) AND Concept 3: TX (“clinical trial*” OR outcome*)	
Social Science Abstracts	Peer-reviewed	Concept 1: DE “suicide” OR DE “Attempted Suicide” OR DE “Suicidal Ideation” AND Concept 2: DE “clinical trials” OR DE “treatment outcomes”	72 articles After removing duplicates, non-English language articles, and off-target reports, 26 articles were retained.
	Peer-reviewed; free text search	Concept 1: TI (suicid*) AND Concept 2: TX (prevent*) AND Concept 3: TX (“clinical trial*” OR outcome*)	
Defense Technical Information Center	Technical reports	Concept 1: TI:(suicid*) AND Concept 2: (“clinical trial*” OR outcome*) AND Concept 3: prevent*	54 articles After removing duplicates and off-target reports, 5 articles were retained.

NOTE: An asterisk (*) denotes a wildcard search. MeSH is an abbreviation for medical subject heading. TI is an abbreviation for title, SU for subject, DE for descriptor, and TX for anywhere in the text.

Table 2.2
Inclusion and Exclusion Criteria

Inclusion Criteria		Exclusion Criteria
First-tier-specific	Included articles that described the evaluation of a program whose primary goal is to reduce or prevent suicide Included articles that described the results of an empirical study	Articles that described the use of a suicide prevention/reduction program outside the realm of suicide prevention and reduction (e.g., an evaluation of an SPP being used to treat drug addiction) Articles that described a program or plans for an assessment without providing an empirical evaluation Articles that focused on the epidemiology of suicide Articles published in a language other than English Editorials, letters, and commentaries
Second-tier-specific	Included articles that focused on the methodology of evaluating suicide prevention/reduction programs Included articles that reviewed the literature on the evaluation of SPPs	
General	Included articles that employed a range of experimental strategies Included articles from different countries	

program or any information relevant to the evaluation of SPPs, which frequently was not revealed in the title or abstract. An additional nine articles were found during the full-text review in the references of other articles reviewed and were added for inclusion. All remaining articles (n = 200) were included in either the first-tier (n = 166)—which included articles from phase 1 (n = 44) and phase 2 (n = 122) searches—or the second-tier (n = 34) review, which included articles from phase 2 searches only. Figure 2.1 shows a flowchart of the identification and exclusion process. For a full list of articles reviewed, see the references section at the end of the toolkit.

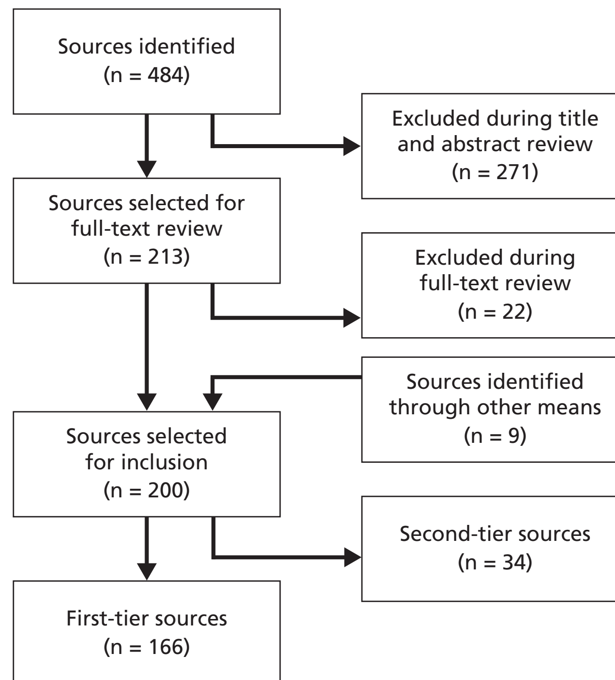
Abstracting Consistent Information from Each Source

Article Coding and Data Abstraction

For each included first-tier article, we abstracted two sets of information: evaluation data and measure data. Evaluation data included a description of the suicide prevention/reduction program, details about the evaluation design, and a synopsis of the study's findings. Measure data described how the study outcomes were assessed and details about measure administration, scoring, and reliability. Each piece of abstracted information represents a characteristic or quality of a program or measure that was useful to consider when constructing the toolkit. Abstracted information was categorized as follows:

- Evaluation data:
  – *Program Description.* This included the name of the program, a description of the targeted population (e.g., specific age group, military members), and a description of the program (including duration and purpose).

Figure 2.1
Flowchart for Literature Search



RAND *RR283-2.1*

- *Evaluation Design.* This included a description of the design and details about data collection time points, sample sizes, and the use of randomization and control groups.
- *Evaluation Findings.* This included a summary of any significant findings reported in the article.

- Measure data:
  - *Measure Description.* This included the measure name, measure category (a description of what was being measured), and the original source of the measure.
  - *Measure Administration.* This included the method of administration (e.g., questionnaire, interview), the person who administered the measure (e.g., self, trained clinician), and the frequency of administration.
  - *Measure Scoring.* This included information about subscales, response options, the presence or absence of a clinical cutoff score, and sample items.
  - *Psychometric Properties.* Recorded information included assessments of reliability and validity.

Tables 2.3 and 2.4 provide further details about each of these abstracted pieces of information.

Table 2.3
Evaluation Data Abstraction Form

Elements Abstracted from Each Article	Brief Description of Each Element
Reference	The full reference information for the article
Program Description	
Program name	The full name of the program, if available
Targeted population	A description of the population targeted by the program (e.g., age group, military/nonmilitary, location)
Brief description of program	A one- to two-sentence description of the program, including duration and purpose
Evaluation Design	
Brief description of evaluation design	A one- to two-sentence description of the evaluation design, including an indication of whether the study used randomization and comparisons over time or between conditions
Time points for data collection	A description of the frequency and time intervals of data collection in the evaluation (e.g., longitudinal, cross-sectional, single time point)
Sample size (n)	The total number of subjects in the sample, as well as the sample size for each condition
Whether the evaluation was experimental	Whether the evaluation was experimental, quasi-experimental, or neither (i.e., Was there a control or comparison group?)
Evaluation Findings	
Brief description of evaluation findings	A one- to two-sentence description of the evaluation findings, including the presence or absence of any statistically significant results

Table 2.4
Measure Data Abstraction Form

Elements Abstracted from Each Article	Brief Description of Each Element
Reference	The full reference information for the article addressing the measure
Program name	The full name of the program being measured, if available
Measure Description	
Measure category	A categorical description of the measure: Knowledge about suicide Knowledge about suicide or mental health–related resources Attitudes about suicide Attitudes toward mental health treatment Skills associated with help-seeking behaviors Problem-solving skills Suicide intervention skills Self-efficacy in identifying and referring individuals at risk Screening for self-damaging, impulsive behavior Screening for suicide risk Screening for mental health and substance abuse problems Dosage Patient satisfaction Fidelity Presence or severity of mental health problems Barriers to care Treatment adherence Medication use Emotional state Means-restriction education Development of a plan for restriction of means among individuals at risk for suicide Firearm access Other restriction of means (measures that capture whether/how much means were restricted) Suicidal ideations Suicide attempts Suicide (i.e., death) Substance misuse, abuse, and dependence Psychological correlates (e.g., hopelessness, impulsivity, problem-solving deficits) Genetics or neurobiology Child abuse "Triggering events" (e.g., relationship problem, financial problems)
Measure name[a]	The full or official name of the measure
Measure Administration	
Method of administration	A description of how the measure was administered (e.g., questionnaire, interview, observational)
Who administered the measure	A description of who administered the measure (e.g., self-administered, clinician, trained professional)
Frequency of administration	The number of times the measure was administered
Measure Scoring	
Clinical cutoff score	Whether the measure had a clinical cutoff score and what the clinical cutoff score was

Table 2.4—Continued

Elements Abstracted from Each Article	Brief Description of Each Element
Measure Scoring (cont.)	
Measure subscales	Whether the measure had subscales and a list and brief description of each, including the number of items per subscale (A subscale is a group of individual items that when considered together provide information about the same characteristic.)
Response options and anchors	One to two sentences describing the response options and anchors for each scale in the measure (e.g., 1 [not at all] to 7 [very much]); response options and anchors for each scale if the evaluation had multiple scales
Sample item	A representative measure item, if available, and the total number of items on the measure
Reliability and Validity of the Measure	
Psychometrics	A description of the reliability and validity of the measure (Reliability was commonly reported as Cronbach's alpha, test-retest reliability, or internal consistency.)
Locating the Measure	
How to obtain the measure	A description of how to obtain a copy of the measure or the reference(s) in which the measure was published

[a] Some measures did not have official names because they were developed for a single study or a specific program.

We used the second-tier articles to supplement existing information in the toolkit chapters. Because information in these articles varied widely, we could not use a standard abstraction or coding procedure.

Procedure for Coding Articles

Coders received initial instruction on the use of the data abstraction form and the content to be included. Each then coded two articles, which were reviewed and discussed by the team. The remaining articles were distributed among the team for independent coding. The team regularly reviewed questions about coding to ensure reliability and consistency among the coders. When all articles had been coded, the first author reviewed the data abstraction forms for completeness and clarity.

Review of Other Evaluation Toolkits

Identifying the Key Components of Evaluation Toolkits

To help inform the drafting of our toolkit, we first reviewed existing evaluation toolkits to develop an outline. We relied, in particular, on the Getting To Outcomes® (GTO) approach because GTO is the only evidence-based model and intervention proven to increase programs' ability to conduct self-evaluation (see Acosta and Chinman, 2011, and two forthcoming articles by Chinman et al.). GTO is based on the theories of traditional evaluation, empowerment evaluation, results-based accountability, and continuous quality improvement. If implemented, the ten steps of GTO will help programs achieve results and demonstrate program

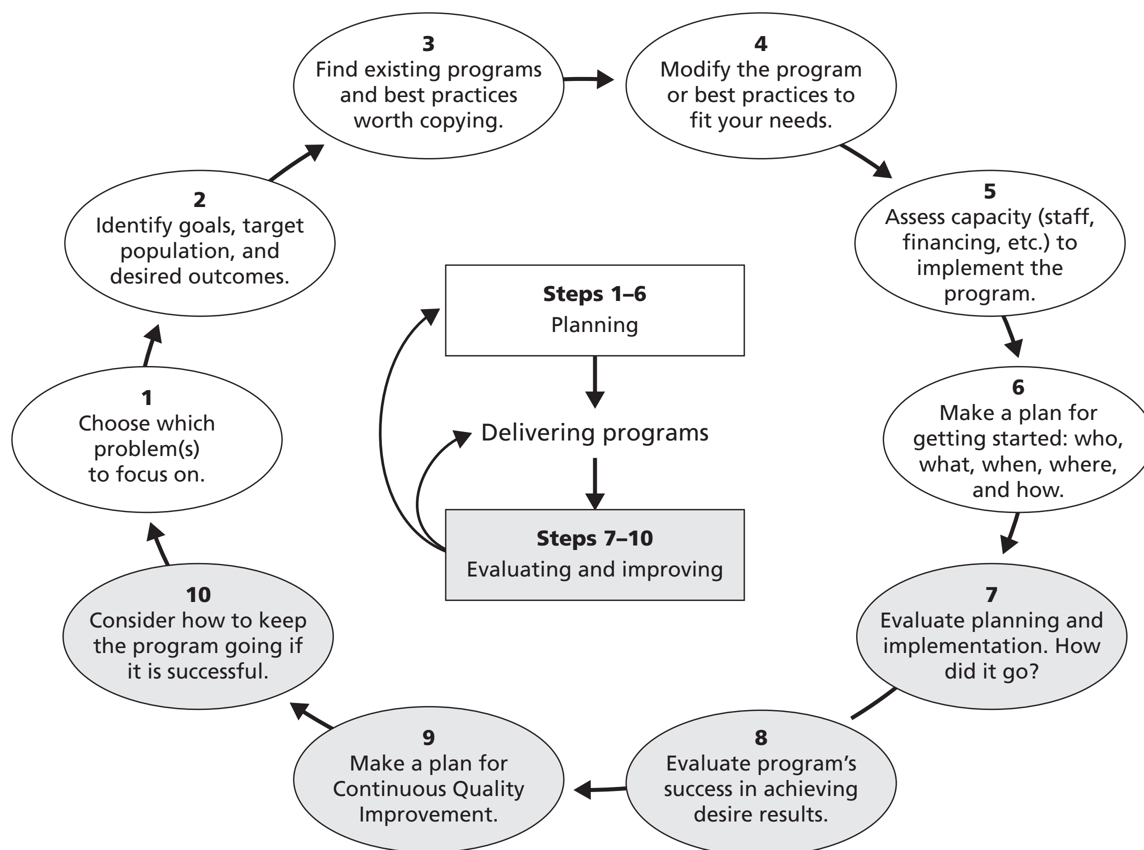
accountability to key stakeholders (e.g., funders). Figure 2.2 shows the ten steps of GTO. It is important to note that these steps are typical program planning activities and that are often not undertaken in a linear fashion.

In addition to the GTO manuals, we also reviewed the following resources:

- the Suicide Prevention Action Network booklet *Suicide Prevention: Effectiveness and Evaluation* (2001)
- *Assessment and Planning Tool Kit for Suicide Prevention in First Nations Communities,* prepared for the First Nations Centre of Canada's National Aboriginal Health Organization (2005)
- an evaluation handbook developed by the W. K. Kellogg Foundation (2000).

We reviewed these toolkits to determine the types of tools that they offered. Tools identified included sample measures, checklists with yes/no questions, worksheets with open-ended questions that help guide users through the toolkit, and designs of prior evaluation studies. Table 2.5 summarizes our review of these sources.

Figure 2.2
Getting To Outcomes Ten Steps to High-Quality Prevention



SOURCE: Chinman et al., 2008.
RAND *RR283-2.2*

Table 2.5
Review of Tools Offered in Existing Toolkits

Evaluation Toolkit	Sample Measures	Checklists	Worksheets	Designs of Prior Evaluations
GTO	X	X	X	Limited to logic model examples
Evaluation toolkit developed by the W. K. Kellogg Foundation		X	X	Limited to logic model examples
Suicide Prevention Action Network booklet on suicide		X		None
Assessment and Planning Tool Kit for Suicide Prevention in First Nations Communities		X		None

Outlining the Chapters of the Toolkit

As mentioned earlier, we relied most heavily on GTO for our toolkit outline, focusing primarily on GTO step 1 (choosing a problem), which includes the development of a logic model, step 7 (process evaluation), step 8 (outcome evaluation), step 9 (continuous quality improvement, or CQI), and step 10 (sustainability). We designed our toolkit to help individuals currently implementing a SPP as they design a program evaluation approach. We organized the toolkit into six chapters, with five chapters focused specifically on the GTO steps (Chapters Two through Six).

- Chapter One introduces users to the toolkit, provides an overview of its contents, and summarizes the methods used to develop the toolkit. The objective of this chapter is to explain the purpose and content of the toolkit and help users decide whether the toolkit is appropriate for use with their program.
- Chapter Two helps users identify the core components of their SPP and organize them into a logic model to clearly visualize the relationships and dependencies between components. This chapter also contains tools to help users review their logic model, assessing whether it is complete and reasonable. The logic model is based on templates offered by GTO and the evaluation toolkit developed by the W. K. Kellogg Foundation.
- Chapter Three helps users select an evaluation design by sharing information from prior evaluations.
- Chapter Four helps users select process and outcome evaluation measures by sharing information from prior evaluations.
- Chapter Five describes how to enter and analyze the evaluation data that users have collected. It includes step-by-step instructions on conducting some basic data analysis using Microsoft Excel® 2010.
- Chapter Six helps programs apply their process and outcome evaluation data to guide improvements. The chapter also walks users through GTO steps 1–6 to help them determine areas in need of improvement.

- The toolkit also contains an appendix with short summaries of each of the first-tier articles we reviewed to inform the development of the toolkit. A second appendix offers a glossary of terms used in the toolkit.

The first-tier articles contained information about evaluation design, measures and associated analyses, and anticipated outcomes for each type of SPP. These articles were used to inform Chapters Three through Six of the toolkit, which focus on evaluation design, measures, and data analysis, respectively. Each of the first-tier articles is also summarized briefly in Appendix A of the toolkit. This allows users to learn more about prior evaluations of programs similar to their own SPP. The second-tier articles contained information about methodological challenges to evaluating suicide, including specific challenges with regard to evaluation design, measurement, and data analysis, as well as reflections on the general state of suicide measurement. Similar to the first-tier articles, information from these second-tier articles was also used to inform Chapters Three through Six, as well as Chapter One, which provides a brief overview of the evaluation challenges associated with the study of suicide. For Chapter One, we also relied on information from an earlier RAND study on suicide prevention, which recommended the development of this toolkit (Ramchand et al., 2011). Table 2.6 provides a summary of the resources that informed the toolkit's chapters.

Table 2.6
Summary of Sources Used to Inform the Toolkit's Development, by Chapter

Toolkit Chapter	Sources
Chapter One. Introduction and Overview	Second-tier literature Ramchand et al., 2011 This report
Chapter Two. Identify Your Program's Core Components and Build a Logic Model	GTO manuals Evaluation handbook developed by the W. K. Kellogg Foundation
Chapter Three. Design an Evaluation for Your Program	GTO manuals First- and second-tier literature
Chapter Four. Select Evaluation Measures for Your Program	GTO manuals First- and second-tier literature
Chapter Five. Analyze Your Program's Evaluation Data	GTO manuals First- and second-tier literature
Chapter Six. Use Evaluation Data to Improve Your Program	GTO manuals First- and second-tier literature
Appendix A. Summary of Evaluation Studies, by Program Type	First-tier literature

NOTE: This table reflects the chapter organization used in the current version of the toolkit.

CHAPTER THREE

Review and Feedback on the Toolkit by DoD Suicide Prevention Program Staff

Once an initial draft of the toolkit was developed, we shared the draft with 12 program staff responsible for implementing SPPs in the Air Force, Marine Corps, and National Guard and solicited their feedback. This chapter describes the process used to engage the program staff and elicit feedback and summarizes the feedback we received. This chapter ends by describing revisions made to the toolkit as a result of feedback from program staff.

Identifying Program Staff to Review the Toolkit

RAND researchers attended a meeting of the Suicide Prevention and Risk Reduction Committee (SPARRC), which oversees DoD suicide prevention efforts, to present on the toolkit and ask for help in recruiting program staff to review it. SPARRC members then sent emails to individual program staff informing them of the opportunity to review the toolkit.

Seventeen program staff representing 13 SPPs expressed an interest in participating in the review. A RAND researcher followed up to help determine their level of interest, explain the procedures and timeline for review of the toolkit, and identify a physical location to mail their copy of the toolkit. Of the 17 individuals who initially expressed interest, 12 program staff, each representing a different program, were available to participate in the review. Programs included service member and leader trainings to raise awareness of the signs and symptoms of suicide (n = 5), gatekeeper trainings (n = 4), screening programs to identify service members at increased risk for suicide (n = 1), behavioral health provider training on suicide risk assessment and management (n = 1), and the Yellow Ribbon program (n = 1). Programs ranged from those focusing on a single installation to those that operated in multiple installations. Participating staff included four civilians, two majors, two first lieutenants, two lieutenant commanders, one physician, and one psychiatrist.

Eliciting Feedback from Program Staff

Program staff were asked to spend five to six hours reading through the toolkit over the course of a month and completing the worksheets, checklists, and templates. Program staff were asked to compile their feedback using a standardized feedback form (see the appendix to this report) as they reviewed the toolkit. The feedback form asked staff to rate the extent to which each chapter of the toolkit met its objectives, whether there were any sections of the toolkit that were not clear or that were difficult to understand, and whether program staff felt uncomfortable

using any of the tools provided. Program staff emailed completed feedback forms to RAND researchers and then participated in a follow-up conference call to discuss their feedback.

Of the 12 program staff who committed to reviewing the toolkit, we received completed feedback forms from ten. One participant preferred to provide feedback through a phone call with a RAND researcher, and one was unable to complete the review for personal reasons.

As a thank you for the time spent on the review, the participating program staff were provided up to four hours of in-kind consultation related to the evaluation of their programs.

It is important to note that we did not collect completed worksheets and templates from pilot test participants. Therefore, we were not able to do a quality assessment on participants' final evaluation designs. In addition, program staff feedback was self-reported. We were unable to observe program staff interacting with the toolkit, which limited our ability to objectively assess comprehension and application of the material.

Program Staff Feedback

During their review, program staff used some or all of the chapters of the toolkit (see Table 3.1). All program staff reviewed Chapters One through Four. Program staff who had access to evaluation data also reviewed Chapters Five and Six.

Program staff were asked to rate the extent to which each chapter met its defined objectives on a scale ranging from strongly agree (5) to strongly disagree (1). Overall, staff agreed or strongly agreed that the toolkit chapters met each of their objectives. Specific chapter objectives and program ratings are shown in Table 3.2.

Table 3.1
Toolkit Chapters Reviewed by Program Staff

	Program Staff									
Toolkit Chapter	1	2	3	4	5	6	7	8	9	10
Chapter One. Introduction and Overview	X	X	X	X	X	X	X	X	X	X
Chapter Two. How the Toolkit Was Developed[a]	X	X	X	X	X	X	X	X	X	X
Chapter Three. Identify Your Program's Core Components for Evaluation and Build a Program Logic Model	X	X	X	X	X	X	X	X	X	X
Chapter Four. Design an Evaluation for Your Program	X	X	X	X	X	X	X	X	X	X
Chapter Five. Select Evaluation Measures for Your Program	X	X	X	X	X	X	X	X	X	X
Chapter Six. Analyze Your Program's Evaluation Data			X	X		X	X		X	X
Chapter Seven. Use Evaluation Data To Improve Your Program			X	X		X	X		X	X

[a] After the pilot test, the "Introduction and Overview" and "How the Toolkit Was Developed" chapters were combined into a single chapter; program staff rated these chapters separately during the pilot test.

Table 3.2
Extent to Which Toolkit Chapters Met Their Objectives

Toolkit Chapter Objectives	Median, Range
Chapter One. Introduction and Overview	
The chapter clearly explained the purpose and content in the toolkit.	5, 4–5
The chapter helped me decide whether this toolkit is appropriate for use with my program.	5, 4–5
Chapter Two. How the Toolkit Was Developed[a]	
The chapter briefly summarized how the toolkit was developed.	5, 4–5
The chapter directed me to Appendix A for more detail on the methods.	5, 4–5
Chapter Three. Identify Your Program's Core Components for Evaluation and Develop a Program Logic Model	
The chapter helped me identify the core components of my program.	5, 4–5
The chapter provided guidance on how to develop a logic model.	5, 4–5
Chapter Four. Design an Evaluation for Your Program	
The chapter provided guidance about the type of evaluation appropriate for my program.	5, 3–5
The chapter helped me select an evaluation based on the resources and expertise my program has available.	4, 3–5
Chapter Five. Select Evaluation Measures for Your Program	
The chapter helped me select process evaluation measures.	4, 3–5
The chapter helped me select outcome evaluation measures.	4, 2–5
Chapter Six. Analyze Your Program's Evaluation Data	
The chapter described how to enter evaluation data into a database.	4, 3–5
The chapter described how to analyze evaluation data.	4, 3–5
Chapter Seven. Use Evaluation Data to Improve Your Program	
The chapter described how to use my evaluation data for program improvement.	5, 4–5

NOTE: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree.
[a] After the pilot test, the "Introduction and Overview" and "How the Toolkit Was Developed" chapters were combined into a single chapter; program staff rated these chapters separately during the pilot test.

Program staff were also asked to indicate (yes/no) whether each chapter was clear and easy to understand and whether they had any difficulty using the tools in the chapter (see Table 3.3). Overall, most program staff felt that the chapters were clear and the tools were easy to use. The only exception was in Chapter Five, which provides step-by-step instructions to help program staff analyze evaluation data using Microsoft Excel. Program staff struggled the most with this chapter, with 50 percent indicating that some sections in the chapter were unclear or difficult to understand and that they had some difficulty using the primers. In the debriefing conference call, a RAND researcher asked program staff for specific suggestions for improvement, and program staff agreed that the best way to improve this chapter would be to automate the data analysis by developing an interactive website that would allow users to view videos of data analysis strategies. This would provide a more interactive guide for data analysis, which

Table 3.3
Extent to Which Chapter Sections Were Clear and Tools Were Easy to Use

	Program Staff (%)	
Toolkit Chapter	**All Sections Were Clear**	**All Tools Were Easy to Use**
Chapter One. Introduction and Overview	100	100
Chapter Two. How the Toolkit Was Developed[a]	100	NA[b]
Chapter Three. Identify Your Program's Core Components for Evaluation and Build a Program Logic Model	100	90
Chapter Four. Design an Evaluation for Your Program	100	90
Chapter Five. Select Evaluation Measures for Your Program	90	90
Chapter Six. Analyze Your Program's Evaluation Data	50	50
Chapter Seven. Use Evaluation Data to Improve Your Program	100	100

[a] After the pilot test, the "Introduction and Overview" and "How the Toolkit Was Developed" chapters were combined into a single chapter; program staff rated these chapters separately during the pilot test.
[b] There were no tools to be rated in Chapter Two.

program staff felt would be easier to follow than the screenshots used in the chapter. Unfortunately, creating this interaction guide was outside the scope of the current project.

Program staff were also provided space to offer additional feedback on how to improve the toolkit. Four participants suggested that the toolkit chapters, particularly Chapters Five and Six, would be easier and more efficient to use as an online program. Since many of the toolkit worksheets, templates, and checklists are linked, program staff must transfer information from worksheets to templates in the toolkit. A programmed online version of the toolkit would allow for an autofill feature to prepopulate this information and would lower the burden on program staff, who currently have to transfer this information by hand.

Program staff also brought up two issues that were not about the content or design of the toolkit, but larger service-specific issues. Two participants mentioned that it was difficult to find the time to conduct an evaluation and really work through the steps outlined in the toolkit. Two participants also questioned how best to merge these suicide prevention efforts with the increasing focus on resilience.

Toolkit Revisions Based on Program Staff Feedback

Based on program staff feedback, the RAND team made several improvements to the toolkit. First, the process evaluation measures table was relocated to Chapter Four. Previously, this table had been an appendix to the toolkit. Second, a footnote was added to the outcome evaluation measures tables to distinguish between free and fee-based measures. Program staff indicated that this was an important consideration for decisionmaking in a resource-limited environment. Finally, a column was added to the process evaluation measures table to provide users with the name and email address of the measure developer so they could request additional information about these measures, if needed. Many of the process evaluation measures

are not readily available online. Therefore, program staff suggested including some guidance in the toolkit about how to get additional information about these measures, if needed.

Although several program staff suggested that the toolkit be converted into a more interactive online product, this conversion was outside the scope of the current project. Therefore, no revisions of this nature were made to the toolkit.

CHAPTER FOUR

Recommendations for Disseminating the Toolkit

A copy of the final toolkit is available as a companion document to this report and can be found at http://www.rand.org/pubs/tools/TL111.html. The toolkit includes worksheets, templates, checklists, and other tools to engage program staff. The RAND researchers responsible for developing the toolkit put together several additional recommendations for DCoE and the Defense Suicide Prevention Office to consider as plans are developed for disseminating the toolkit.

Recommendations

Continue to Refine the Toolkit

As we developed the toolkit, we sought to ensure that it was user-friendly and useful to those who would be encouraged to implement it. We received valuable feedback from a group of program staff responsible for implementing SPPs in the Air Force, Marine Corps, and National Guard. DCoE should develop a plan to regularly review the toolkit content to ensure that it remains updated and relevant to users. For example, as the services develop specific resources for SPPs, these resources could be included in the toolkit. This regular review would allow for additional feedback to be incorporated on a regular basis and optimize the user-friendliness of the toolkit design.

Consider Converting the Toolkit to an Interactive Online Format

As previously mentioned, many of the toolkit worksheets, templates, and checklists are linked and require program staff to transfer information from a worksheet to a template (for example). To lower the burden on users, program staff suggested converting the paper-based toolkit to an interactive online format. This format would allow for an autofill feature to prepopulate information, making it unnecessary for program staff to transfer information by hand. Additionally, an interactive, online manual may be easier to update and keep current.

Continue Partnering with the SPARRC and Defense Suicide Prevention Office to Disseminate the Toolkit

To optimize dissemination of the toolkit at the installation level, SPARRC members, including SPP managers for each branch, should identify a set of dissemination strategies to share the toolkit broadly with potential users across DoD. This could entail hosting webinars to introduce the toolkit, passing out copies of the toolkit at conferences that suicide prevention coordinators may attend, or sending emails to installation-level contacts with an electronic

copy of the toolkit. It is also important to note that interventions to build evaluation capacity, like GTO, have been tested only after practitioners have received formal training and ongoing in-person technical assistance on the use of the manual or toolkit. Therefore, the ease of evaluation may be improved by supplementing the toolkit with training and technical assistance, which could be provided in conjunction with dissemination efforts. This will better meet the varying needs and levels of expertise of program staff.

Conclusion

RAND's suicide prevention evaluation toolkit was developed in response to a recommendation in the 2011 RAND report *The War Within: Preventing Suicide in the U.S. Military* (Ramchand et al., 2011) to evaluate existing programs and ensure that new programs include an evaluation component when they are implemented. The toolkit is intended to build the knowledge and skills of individuals responsible for implementing DoD-sponsored SPPs as they self-evaluate their programs. The content of the toolkit is based on current peer-reviewed literature and feedback from program staff. Dissemination of the toolkit across DoD will help ensure ongoing and future SPPs have some guidance to determine whether their programs produce beneficial effects and will ultimately guide the responsible allocation of scarce resources.

APPENDIX

Feedback Form Used in the Pilot Test of the RAND Evaluation Toolkit for Suicide Prevention Programs

Instructions

We have created a feedback form for each chapter of the toolkit. As you finish reading each chapter and working through tools in each chapter, please pause to answer the corresponding pilot test questions for that chapter. We are looking for your honest opinions, so please answer all the questions.

How do I know if I'm done with the pilot test?

Before sending your feedback back to us, please review your work.

Which chapters of the toolkit did you use?

- ☐ Chapter One. Introduction and Overview
- ☐ Chapter Two. How the Toolkit Was Developed*
- ☐ Chapter Three. Identify Your Program's Core Components for Evaluation and Build a Program Logic Model
- ☐ Chapter Four. Design an Evaluation for Your Program
- ☐ Chapter Five. Select Evaluation Measures for Your Program
- ☐ Chapter Six. Analyze Your Program's Evaluation Data
- ☐ Chapter Seven. Use Evaluation Data to Improve Your Program

** After the pilot test, the "Introduction and Overview" and "How the Toolkit Was Developed" chapters were combined into a single chapter; program staff rated these chapters separately during the pilot test.*

Do you have completed feedback forms for each chapter you selected above?

- ☐ Yes. Great job! You are finished.
- ☐ No. We really want your feedback. Please complete the feedback forms for all chapters of the toolkit that you used during the pilot test.

What if I do not work through all the chapters?

Depending on your program's interest, you may only use a portion of the toolkit. We would appreciate your feedback on any portions of the toolkit you are able to review.

What should I do with my pilot test feedback form?

Please email your feedback back to Joie Acosta at Joie_Acosta@rand.org. She will send an email confirmation letting you know that your feedback has been received. If you prefer to send your feedback by fax, please contact her at 703-413-1100, ext. 5324, to arrange a time for faxing.

What if I have other questions during the pilot test?

Feel free to contact Joie Acosta with any questions you might have at 703-413-1100, ext. 5324, or at Joie_Acosta@rand.org.

Thank you for your feedback!

Chapter One. Introduction and Overview

1. Please indicate the **extent** to which this chapter met its **objectives**.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
The chapter clearly explained the purpose and content in the toolkit.					
The chapter helped me decide whether this toolkit is appropriate for use with my program.					

2. Were there any **sections of the chapter** that were **not clear or difficult to understand**?

☐ No, all sections were clear.
☐ Yes. Which section(s)? Please list page numbers. ____________________

How can we **improve** these sections? (Feel free to write directly on the toolkit.)

3. Did you feel **uncomfortable** using any of the following tools?

☐ No, I felt comfortable completing all the tools in this chapter.
☐ Yes, I had difficulty with the following tools:

☐ Checklist 1.1. Is This Toolkit Right for My Program?
☐ Other (please specify): ____________________

4. Do you have any **additional comments** about how to improve this chapter of the toolkit?

Chapter Two. How the Toolkit Was Developed*

1. Please indicate the **extent** to which this chapter met its **objectives**.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
The chapter briefly summarized how the toolkit was developed.					
The chapter directed me to Appendix A for more detail on the methods.					

2. Were there any **sections of the chapter** that were **not clear or difficult to understand**?

☐ No, all sections were clear.
☐ Yes. Which section(s)? Please list page numbers. ____________________________________

How can we **improve** these sections? (Feel free to write directly on the toolkit.)

____

____

____

____

3. Do you have any **additional comments** about how to improve this chapter of the toolkit?

____

____

____

____

** After the pilot test, the "Introduction and Overview" and "How the Toolkit Was Developed" chapters were combined into a single chapter; program staff rated these chapters separately during the pilot test.*

Chapter Three. Identify Your Program's Core Components for Evaluation and Build a Program Logic Model

1. Please indicate the **extent** to which this chapter met its **objectives**.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
The chapter helped me identify the core components of my program.					
The chapter provided guidance on how to develop a logic model.					

2. Were there any **sections of the chapter** that were **not clear or difficult to understand**?

☐ No, all sections were clear.
☐ Yes. Which section(s)? Please list page numbers. ____________________

⇨ How can we **improve** these sections? (Feel free to write directly on the toolkit.)

____
____
____
____

3. Did you feel **uncomfortable** using any of the following tools?

☐ No, I felt comfortable completing all the tools in this chapter.
☐ Yes, I had difficulty with the following tools:

- ☐ Checklist 3.1. Is Your Logic Model Complete and Appropriately Detailed?
- ☐ Checklist 3.2. Are the Core Components of Your Logic Model Appropriately Aligned?
- ☐ Worksheet 3.1. Identifying Components
- ☐ Templates 3.1 and 3.2. Program Logic Model
- ☐ Other (please specify): ____________________

____

4. Do you have any **additional comments** about how to improve this chapter of the toolkit?

____
____
____
____

Chapter Four. Design an Evaluation for Your Program

1. Please indicate the **extent** to which this chapter met its **objectives**.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
The chapter provided guidance about the type of evaluation appropriate for my program.					
The chapter helped me select an evaluation based on the resources and expertise my program has available.					

2. Were there any **sections of the chapter** that were **not clear or difficult to understand**?

☐ No, all sections were clear.
☐ Yes. Which section(s)? Please list page numbers. ____________________

How can we **improve** these sections? (Feel free to write directly on the toolkit.)

____
____
____
____

3. Did you feel **uncomfortable** using any of the following tools?

☐ No, I felt comfortable completing all the tools in this chapter.
☐ Yes, I had difficulty with the following tools:

- ☐ Checklist 4.1. Which Type of Program(s) Best Describes Your SPP?
- ☐ Appendix B. Summary of Evaluation Studies, by Program Type
- ☐ Checklist 4.2. Does Your Evaluation Plan Consider the Following Factors?
- ☐ Table 4.1. Types of Evaluation Designs
- ☐ Table 4.2. Effects of Suicide Prevention Programs, by Program Type
- ☐ Worksheet 4.1. Issues to Consider for My Program
- ☐ Template 4.1. Evaluation Planner
- ☐ Other (please specify): ____________________

____

4. Do you have any **additional comments** about how to improve this chapter of the toolkit?

____
____
____
____

Chapter Five. Select Evaluation Measures for Your Program

1. Please indicate the **extent** to which this chapter met its **objectives**.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
The chapter helped me select **process** evaluation measures.					
The chapter helped me select **outcome** evaluation measures.					

2. Were there any **sections of the chapter** that were **not clear or difficult to understand**?

☐ No, all sections were clear.
☐ Yes. Which section(s)? Please list page numbers. ____________________

How can we **improve** these sections? (Feel free to write directly on the toolkit.)

____

____

____

____

3. Did you feel **uncomfortable** using any of the following tools?

☐ No, I felt comfortable completing all the tools in this chapter.
☐ Yes, I had difficulty with the following tools:

- ☐ Checklist 5.1. To What Extent Do the Measures Selected Align with Your Program's Target Population, Activities, and Outcomes?
- ☐ Table 5.1. Sample Outcome Measures
- ☐ Appendix C. Sample Process Measures
- ☐ Other (please specify): ____________________

  ____

4. Do you have any **additional comments** about how to improve this chapter of the toolkit?

____

____

____

____

Chapter Six. Analyze Your Program's Evaluation Data

1. Please indicate the **extent** to which this chapter met its **objectives**.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
The chapter described how to enter evaluation data into a database.					
The chapter described how to analyze evaluation data.					

2. Were there any **sections of the chapter** that were **not clear or difficult to understand**?

☐ No, all sections were clear.
☐ Yes. Which section(s)? Please list page numbers. ____________________

⇨ How can we **improve** these sections? (Feel free to write directly on the toolkit.)

____

____

____

____

3. Did you feel **uncomfortable** using any of the following tools?

☐ No, I felt comfortable completing all the tools in this chapter.
☐ Yes, I had difficulty with the following tools:

- ☐ Primer 1: Calculating Descriptive Statistics for Your Program
- ☐ Primer 2: Statistical Models for Detecting Differences in Your Program's Target Population
- ☐ Primer 3: Linking Process to Outcome Measures
- ☐ Other (please specify): ____________________

  ____

4. Do you have any **additional comments** about how to improve this chapter of the toolkit?

____

____

____

____

Chapter Seven. Use Evaluation Data to Improve Your Program

1. Please indicate the **extent** to which this chapter met its **objectives**.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
The chapter described how to use my evaluation data for program improvement.					

2. Were there any **sections of the chapter** that were **not clear or difficult to understand**?

☐ No, all sections were clear.
☐ Yes. Which section(s)? Please list page numbers. ______________________

How can we **improve** these sections? (Feel free to write directly on the toolkit.)

____
____
____
____

3. Did you feel **uncomfortable** using any of the following tools?

☐ No, I felt comfortable completing all the tools in this chapter.
☐ Yes, I had difficulty with the following tools:

- ☐ Checklist 7.1. What CQI Actions Are Needed to Improve the Program?
- ☐ Worksheet 7.1. Description of the Program's Evaluation
- ☐ Worksheet 7.2. Review Program Outcomes
- ☐ Worksheet 7.3. Program Improvement Plan
- ☐ Table 7.1. Results-Based Scenarios and Associated Strategies for Program Improvement
- ☐ Other (please specify): ______________________

____

4. Do you have any **additional comments** about how to improve this chapter of the toolkit?

____
____
____
____

Additional Feedback About the Toolkit

Do you have any **additional comments** about how to improve the toolkit?

References

Acosta, Joie, and Matthew Chinman, "Building Capacity for Outcomes Measurement in Human Service Organizations: The Getting To Outcomes Method," in Jennifer L. Magnabosco and Ronald W. Manderscheid, eds., *Outcomes Measurement in the Human Services: Cross-Cutting Issues and Methods in the Era of Health Reform*, Washington, D.C.: NASW Press, 2011.

Bray, Robert M., Michael R. Pemberton, Laurel L. Hourani, Michael Witt, Kristin L. Rae Olmsted, Janice M. Brown, BeLinda Weimer, Marian E. Lane, Mary Ellen Marsden, Scott Scheffler, Russ Vandermaas-Peeler, Kimberly R. Aspinwall, Erin Anderson, Kathryn Spagnola, Kelly Close, Jennifer L. Gratton, Sara Calvin, and Michael Bradshaw, *2008 Department of Defense Survey of Health Related Behaviors Among Active Duty Military Personnel*, Washington, D.C.: U.S. Department of Defense, 2009.

Chinman, Matthew, and Joie Acosta, "Getting To Outcomes: An Example of How to Monitor and Evaluate the Key Functions of a Prevention Support System," *American Journal of Community Psychology*, forthcoming.

Chinman, Matthew, Joie Acosta, Patricia Ebener, Q Burkhart, Michael Clifford, Maryann Corsello, Tim Duffy, Sarah Hunter, Margaret Jones, Michael Lahti, Patrick S. Malone, Susan Paddock, Andrea Phillips, Susan Savell, Peter C. Scales, and Nancy Tellet-Royce, "Intervening with practitioners to improve the quality of prevention: One year findings from a randomized trial of Assets-Getting To Outcomes," *Journal of Primary Prevention*, forthcoming.

Chinman Matthew, Sarah Hunter, Patricia Ebener, Susan Paddock, Lindsey Stillman, Pamela Imm, and Abraham Wandersman, "The Getting To Outcomes Demonstration and Evaluation: An Illustration of the Prevention Support System," *American Journal of Community Psychology*, Vol. 41, No. 3–4, June 2008, pp. 206–224.

First Nations Centre, National Aboriginal Health Organization, *Assessment and Planning Tool Kit for Suicide Prevention in First Nations Communities*, Ottawa, Ont., 2005. As of March 19, 2013:
http://www.naho.ca/documents/fnc/english/FNC_SuicidePreventionToolkit.pdf

Kinn, Julie T., David D. Luxton, Mark A. Reger, Gregory A. Gahm, Nancy A. Skopp, and Nigel E. Bush, *Department of Defense Suicide Event Report: Calendar Year 2010 Annual Report*, Washington, D.C.: National Center for Telehealth and Technology, U.S. Department of Defense, 2011.

Public Law 110-417, National Defense Authorization Act for Fiscal Year 2009, October 14, 2008.

Ramchand, Rajeev, Joie Acosta, Rachel M. Burns, Lisa H. Jaycox, and Christopher G. Pernin, *The War Within: Preventing Suicide in the U.S. Military*, Santa Monica, Calif., MG-953-OSD, 2011. As of March 19, 2013:
http://www.rand.org/pubs/monographs/MG953.html

Suicide Prevention Action Network, *Suicide Prevention: Prevention Effectiveness and Evaluation*, Washington, D.C., 2001. As of March 19, 2013:
http://www.sprc.org/sites/sprc.org/files/library/prevtoolkit.pdf

U.S. Department of Defense Task Force on the Prevention of Suicide by Members of the Armed Forces, *The Challenge and the Promise: Strengthening the Force, Preventing Suicide, and Saving Lives*, Washington, D.C., August 2010. As of March 19, 2013:
http://www.health.mil/dhb/downloads/Suicide%20Prevention%20Task%20Force%20final%20report%20 8-23-10.pdf

W. K. Kellogg Foundation, *Logic Model Development Guide*, Battle Creek, Mich., October 2000.