U.S. Department of Veterans Affairs
Veterans Health Administration
Quality Enhancement Research Initiative

Quality Enhancement Research Initiative (QUERI) Evaluation Guide

CEIR
VA QUERI Center for Evaluation
and Implementation Resources

QUERI
VA Quality Enhancement Research Initiative
EVIDENCE INTO PRACTICE

Acknowledgements

Suggested Citation

This Department of Veterans Affairs QUERI Evaluation Guide was developed by the Department of Veterans Affairs Quality Enhancement Research Initiative (QUERI) to support evaluations of clinical practices, programs, and initiatives within health care settings. The views expressed in this document do not necessarily reflect those of the Department of Veterans Affairs.

Contributors

Graphic Design and Editorial Assistance

For More Information

Center for Evaluation and Implementation Resources

Quality Enhancement Research Initiative

U.S. Department of Veterans Affairs

ceir@va.gov

Table of Contents

Introduction

The Foundations for Evidence-Based Policymaking Act of 2018 (or "Evidence Act")[1] requires all cabinet level agencies, including the Department of Veterans Affairs (VA), to ensure their budgets and policies are tied to, supported, and justified by evidence. It highlights the essential role of evaluation in helping policymakers determine whether activities have intended outcomes and impacts. The current VA budget process places priority on scale-up interventions or policies that have been shown to work and proposals that will further develop the agency's capacity to use evidence, evaluation, and data as tools to improve program outcomes.

A significant tool to support greater organizational effectiveness and impact is the Learning Health System Model, which uses data and technology to identify and address gaps in care quality through a three-part cycle that supports the integration of data into practice via efforts to transition: (1) data to knowledge; (2) knowledge to performance; and (3) performance to data.[2] As VA transforms into a learning health system,[3] comprehensive evaluation of the effectiveness, value, and impact of programs has become the standard for high-quality health care delivery and policy. Evaluation is a crucial step in determining whether programs and policies were implemented as intended, with an eye towards assessing the timeline for implementation, effectiveness, impact, and ultimately, sustainability. The VA Quality Enhancement Research Initiative (QUERI) Center for Evaluation and Implementation Resources (CEIR) developed this guide to highlight best practices for conducting high-quality evaluations. Designed to support the development, planning, and implementation of rigorous evaluations, the guide provides references, examples, and targeted materials that may be used side-by-side in designing an evaluation.

This guide is designed to be used by persons interested in all phases of evaluation, from how to design and carry out an evaluation, to how interpret the results. Although the information presented in this guide is based on expert best practice recommendations, it is not intended to be a stand-alone resource. Instead, users are encouraged to apply the information in this guide while also seeking outside consultation with persons that have experience in designing and executing evaluations, ideally in the settings in which new evaluations are to be developed.

The **primary audience** for this guide includes VA employees, investigators, and leaders; however, much of the content can also apply to evaluation work conducted by other federal agencies.

Note: For the purposes of this guide, "program" is used to refer to any program, policy, service, intervention, practice, or innovation that would be the focus of an evaluation.

---

[1] Foundations for Evidence-Based Policymaking Act of 2018. Public Law 115-435. 132 STAT 5529 (2018). Retrieved from https://www.congress.gov/bill/115th-congress/house-bill/4174

[2] Guise JM, Savitz LA, Friedman CP. Mind the gap: putting evidence into practice in the era of learning health systems. *J Gen Intern Med. 2018*; 33:2237-9.

[3] Atkins D, Kilbourne AM, Shulkin D. Moving from discovery to system-wide change: The role of research in a Learning Health Care System: Experiences from three decades of health systems research in the Veterans Health Administration. *Annu Rev Public Health*. 2017;38:467–87.

What is Evaluation?

The Evidence Act provides guidance on activities that meet the standards for evaluations within federal agencies as well as expectations for characteristics of a high-quality evaluation. Within the Evidence Act, **program evaluation** is defined as "an assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency." Evaluation asks questions such as "Does the program work?", "What makes it work?", and "How can we make it work in the real world?" Results from high-quality evaluations can highlight aspects of programs or services that could be improved as well as the steps that might be taken to achieve improvement. Evaluations can also inform the development of new policies and directives aimed at improving the quality of care provided to consumers and enhancing the work environment for staff.

Is Evaluation Considered Research?

VA evaluations can be classified as research or non-research. The Office of Research and Development provides a Program Guide with guidance[4] on determining whether a Veterans Health Administration (VHA) operations activity constitutes research and the process for documenting the activity as non-research, if it is determined to be non-research.

Research as defined by Program Guide 1200.21 is "a systematic investigation (including research development, testing, and evaluation) designed to develop or contribute to generalizable knowledge." It may also be defined as a systematic investigation designed to produce information to expand the knowledge base of a scientific discipline (or other scholarly field of study). The Guide also lists activities that are *always* considered research, such as clinical investigations as defined under Food and Drug Administration regulations and those funded or otherwise supported as research by the Office of Research and Development (ORD) or any other sponsor.

As of 2021, activities funded through the VHA Congressional research appropriation—notably projects funded through ORD's research services and programs (e.g., Genomics, Biomedical, Clinical, Health Services, and Rehabilitation Research and Development)—are *almost always* considered research.

VA non-research activities are activities designed and implemented for internal VA purposes with the expectation that: (1) the primary application of findings will be used "by and within VA or by entities responsible for overseeing VA, such as Congress or the Office and Management and Budget"; and (2) the activity is not designed to produce information that expands on the knowledge base of a scientific discipline."

---

[4] Department of Veterans Affairs Office of Research and Development. Program Guide 1200.21. VHA Operations Activities that May Constitute Research. January 9, 2019. Retrieved from https://www.research.va.gov/resources/policies/ProgramGuide-1200-21-VHA-Operations-Activities.pdf

Is My Evaluation Considered Research or Non-Research?

The Common Rule[5] is the 1981 baseline standard of ethics by which any United States government-funded biomedical and behavioral research involving human subjects is held accountable. Recent changes to the Common Rule—including the broadening of research exemption categories—enable certain evaluations to be designated as non-research or "quality improvement." As a result, more program evaluations are being designated as non-research,[6] especially evaluations that assess the impacts of a program on the health system.

VA-Specific Guidance
The following online seminar provides guidance on how to best navigate differences between research and non-research evaluation: Jeans CK, Kilbourne AM, Booker F, Braganza MZ. VA Quality Enhancement Research Initiative (QUERI): Everything you need to know about a QUERI non-research ("QI") protocols. 2019. Available at: https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/3669-notes.pdf

In VA, evaluations designed to inform improvement efforts within the health system can be designated as non-research if the work is not designed to produce information that expands the knowledge base of a scientific discipline or other scholarly field. Note that randomization alone does not define an evaluation as research. Non-research quality improvement initiatives can involve randomization as long as the focus is on internal program evaluation and not testing a new treatment or practice for efficacy or effectiveness on patient health.

Several factors are useful for determining whether a planned activity is research or non-research. Any activity whose conceptualization, plan, or implementation is supplemented or modified in order to produce information expanding the knowledge base of a scientific discipline is considered research. While not always research, nearly any project design that changes the intervention process via techniques such as double-blinding, placebo controls, or prospective consumer-level randomization to clinical interventions that are not tailored to individual consumer benefit are considered research activities. Individuals conducting operations activities have a responsibility to consult their supervisors as soon as possible whenever there may be doubt about the research versus non-research status of an operations activity.

Can I Publish Evaluations Considered Non-Research?

Evaluators can publish non-research projects in scientific journals but must document non-research status for most journals. It is recommended they consult with the journals

---

[5] Office for Human Research Protections. Revised Common Rule 2018. Federal Register Vol 82 No 12. January 19, 2017. Retrieved from https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html
[6] Horwitz LI, Kuznetsova M, Jones SA. Creating a learning health system through rapid-cycle, randomized testing. *N Engl J Med*. 2019;381(12):1175-1179.

and their local health system or academic research office regarding the best approach for documenting non-research status and whether the protocol would need to be reviewed by an Institutional Review Board.

In many cases, evaluations can still be considered "research" but not "human subjects research." In these cases, the evaluations are not reviewed for human subjects protections by an Institutional Review Board. The local health system or academic research office can be consulted for additional information on the distinction of research versus non-research. It is important to remember that in VA all research must be approved by a Research and Development Committee.

Why Evaluate?

Evaluation is an essential tool to support evidence-based policymaking that ensures taxpayer dollars are spent wisely on programs that ultimately benefit populations in need. The Evidence Act requires that federal agency budgets be justified by evidence and require the adoption of evaluation as a core function to ensure effective budget processes over time. Evaluations, especially those that are of high quality and are subject to rigorous peer review (e.g., see How do I find funding for my evaluation? at the end of this guide), enable agencies' maximum confidence in the results. Moreover, agencies that devote portions of their budgets to evaluation activities are recognized by the Office of Management and Budget for supporting evidence-based policy. Finally, results from the evaluation can inform what resources, policies, or infrastructure are required to sustain the program or policy, if demonstrated effective.

Programs are evaluated for several reasons including[7]:

- Addressing questions related to the implementation or institution of a program, policy, or organization.

- Addressing the effectiveness/impact of specific strategies related to or used by a program, policy, or organization; and/or factors that relate to variability in the effectiveness of a program, policy, or organization or strategies of these. This can include assessments of clinical and patient outcomes to determine whether particular interventions are maintaining clinical effectiveness in real-world settings.

- Examining questions related to understanding the contextual factors surrounding a program, as well as how to effectively target specific populations or groups for a particular intervention.

- Providing critical information to inform decisions about current and future programming, policies, and organizational operations.

- Enhancing learning and improvement purposes, as well as accountability purposes.

Within VA and other federal agencies, evaluation takes on an even greater role in light of the Evidence Act. As noted above, the Evidence Act requires all cabinet-level agencies, including VA, to ensure their budgets and policies are tied to, supported by, and justified by evidence. The Evidence Act further stipulates that the Director of the

---

[7] Adapted from Office of Management and Budget recommendations for implementation of the Evidence Act, 3/10/2020

Office of Management and Budget shall conduct biannual evaluations of the extent to which Evidence Act guidelines have been effectively adopted into practice within government agencies.

Types of Evaluation

Evaluations generally fall into one of two broad categories: (1) Formative; or (2) Summative.[8] Within these, there are also different types of evaluations (Figure 1). Which of these evaluations is most appropriate depends on the central evaluation questions and the stage of the program, keeping in mind that an evaluation can inform multiple phases of program development and implementation. Figure 1 presents the continuum of evaluation.

Figure 1: Continuum of Evaluation Types Based on Program Stage and Central Questions[9]

Program Stage	Before Program Begins	New Program	Established Program	Mature Program
Formative → Summative				
Evaluation Type	Needs Assessment	Process/ Implementation Evaluation	Outcome Evaluation	Impact Evaluation
Questions	What is the need? What can be done to address the need?	Is the program operating as planned?	Is the program achieving its objectives?	What predicted and unpredicted impacts has the program had?
	How can we reduce re-admissions for behavioral concerns?	*How often did team members visit the Community Facility?*	*What is the re-admission rate for residents discharged to the community?*	*Were other facilities more likely to accept residents with behavioral concerns?*

Formative evaluations are typically conducted to assess whether a program, policy, or organizational approach—or some aspect of these—is feasible, appropriate, and acceptable before it is fully implemented. They may include process and/or outcome measures. However, unlike outcome and impact evaluations, which seek to answer whether the program, policy, or organization met its intended goals or had the intended impacts, formative evaluation focuses on learning and improvement and does not aim to answer questions of overall effectiveness.[10] Such evaluations are useful for determining

---

[8] Definitions are adapted from the CDC. See Salabarría-Peña, Y, Apt, B.S., Walsh, C.M. Practical Use of Program Evaluation among Sexually Transmitted Disease (STD) Programs, Atlanta (GA.
Centers for Disease Control and Prevention; 2007.

[9] Centers for Disease Control and Prevention. Evaluation planning: What is it and how do you do it? Retrieved from https://www.cdc.gov/healthcommunication/pdf/evaluationplanning.pdf

[10] Office of Management and Budget. Circular No. A-11. Preparation, Submission, and Execution of the Budget. June 2019. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2018/06/a11.pdf

how to achieve program goals and improve new programs. Two common types of formative evaluation include needs assessment and process/implementation evaluation.

- **Needs assessment** is a process for determining an organization's or population's needs—what the gaps in service provision are, and what the desired future state is. Needs assessments aim to answer: *What is the need? Who needs the program? How great is the need? How might the need to be best met?* An example of a needs assessment within VA is Ellison et al. (2012),[11] in which qualitative interviews were used to assess the perceived educational needs of younger (age 18 – 29) and older adult Veterans (age 30+) with self-reported PTSD to understand if age-tailored supported education services were needed.

- **Process evaluation**, also known as implementation evaluation, assesses how the program or service is delivered relative to its intended theory of change, and often includes information on content, quantity, quality, and structure of services provided. Process evaluations answer questions such as: *Were the core components of the program implemented as intended? Was the target population reached? What challenges and barriers were encountered? What strategies were associated with overcoming them? What strategies were associated with successful implementation?* Process/implementation evaluation is also useful to support the dissemination of programs to new sites. Within VA, implementation evaluators often use a combination of methods to identify and understand how multi-level contextual factors at the Veteran-, clinician-, facility-, or health system-level serve as barriers or facilitators to delivery and receipt of quality health care intervention(s) over time. For example, Cohen et al. (2010)[12] utilized surveys and interviews in conducting a process evaluation to identify barriers and facilitators to participation in a family intervention for Veterans with schizophrenia.

Summative evaluations are conducted once programs are more established. They are useful for determining the extent to which a program is achieving its goals. Summative evaluations include outcome evaluation and impact evaluation approaches.

- **Outcome evaluations** measure the extent to which a program, policy, or organization has achieved its intended outcome(s) and focuses on outputs and outcomes to assess effectiveness. Unlike impact evaluations, they typically cannot discern causal attribution. Outcomes can be both short-term and medium-term. Outcome evaluation asks questions such as: *Did the expected changes in outcomes occur? Were program goals met? Can these changes in outcomes be attributed to the program? What components of the program are reproduceable in other locations?* An example of an outcome evaluation within VA is Mattocks et al. (2017),[13] which used both closed- and open-ended survey items to assess Veteran

---

[11] Ellison ML, Mueller L, Smelson D, Corrigan PW, Torres Stone RA, Bokhour BG, Najavits LM, Vessella JM, Drebing C. Supporting the education goals of post-9/11 veterans with self-reported PTSD symptoms: A needs assessment. *Psychiatr Rehab J.* 2012;35(3):209-17.

[12] Cohen AN, Glynn SM, Hamilton AB, Young AS. Implementation of a family intervention for individuals with schizophrenia. *J Gen Intern Med.* 2010;25(1):32-7.

[13] Mattocks KM, Kuzdeba J, Baldor R, Casares J, Lombardini L, Gerber MR. Implementing and evaluating a telephone-based centralized maternity care coordination program for Pregnant Veterans in the Department of Veterans Affairs. *Womens Health Issues.* 2017;27(5):579-85.

satisfaction and use of a maternal care coordination program (MCC), as well as infant and maternal outcomes.

- **Impact evaluations** assess the impact of a program, policy, or organization, or aspect thereof, on outcomes relative to those of a counterfactual. In other words, this type of evaluation estimates and compares outcomes with and without the program, policy, or organization, or aspect thereof. Impact evaluations include both experimental (i.e., randomized controlled trials) and quasi-experimental designs. Impact evaluations ask: *What are the long-term effects of the program—both expected and unexpected?* An example of an impact evaluation within VA is Walsh et al. (1995),[14] which used quantitative and qualitative methods to measure the long-term impact of a four-year series of workshops designed to teach VA employees about the value of teamwork and consultation in caring for geriatric veterans.

Evaluation Frameworks

An evaluation framework is a tool that organizes the essential elements of a program evaluation and identifies steps for effective evaluation. A robust framework can serve as a template for conceiving, planning, and implementing an evaluation. Evaluation frameworks may also help identify factors that should be considered in developing a new program or initiative.

Many frameworks have been developed and the choice of framework is dependent on the type of evaluation being conducted and the questions being asked. The following page details commonly used evaluation frameworks:

---

[14] Walsh PL, Garbs CA, Goodwin M, Wolff EM. An impact evaluation of a VA geriatric team development program. *Gerontol Geriatr Educ.* 1995;15(3):19-35.

Evaluation Framework	Framework Goal	VA Example
Framework for Program Evaluation in Public Health[15]	To help understand program context and improve how program evaluations are conceived and conducted. Provides steps in program evaluation practice and standards for effective program evaluation.	Freeman et al. (2019):[16] Promoting Spiritual Healing by Stress Reduction Through Meditation for Employees at a Veterans Hospital: A CDC Framework-Based Program Evaluation
PRECEDE-PROCEED Model[17,18]	To aid in designing, implementing, and evaluating health promotion and other public health programs. PRECEDE provides the structure for planning a targeted and focused public health program; PROCEED provides the structure for implementing and evaluating the public health program.	Weir et al. (2011):[19] The Role of Information Technology in Translating Educational Interventions into Practice: An Analysis using the PRECEDE/PROCEED Model
RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance[20]	To highlight essential program elements that can improve sustainable adoption and implementation of effective, generalizable, evidence-based programs or interventions.	Etingen et al. (2020):[21] TeleWound Practice Within the Veterans Health Administration: Protocol for a Mixed Methods Program Evaluation
Structure-Process-Outcome Framework (Donabedian Model)[22]	To aid in assessing quality of health care. Structures are the setting for care coordination, including physical or organizational aspects; processes are the modes for care coordination; and outcomes include health outcomes or other measurements.	Watkins et al. (2010):[23] Transforming Mental Healthcare in the Veterans Health Administration: A Model for Measuring Performance to Improve Access, Quality, and Outcomes
Realist Evaluation[24]	To make the theories within a program explicit, specifically by developing clear hypotheses about how, and for whom, programs might 'work'. The evaluation of the program tests those hypotheses.	Hulen et al. (2020):[25] Creating Change, Challenging Structure: Graduate and Faculty Perspectives on the Implementation of an Interprofessional Education Program in Veterans Affairs Primary Care

[15] Centers for Disease Control and Prevention. Framework for program evaluation in public health. MMWR 1999;48(No. RR-11). Retrieved from https://www.cdc.gov/mmwr/PDF/rr/rr4811.pdf

[16] Freeman RC, Sukuan N, Tota NM, Bell SM, Harris AG, Wang HL. Promoting spiritual healing by stress reduction through meditation for employees at a Veterans hospital: A CDC framework–based program evaluation. *Workplace Health Saf.* 2020;68(4):161-70.

[17] Green LW. Toward Cost-Benefit Evaluations of Health Education: Some Concepts, Methods, and Examples. *Health Educ Behav.* 1974;2(1_suppl):34-64.

[18] Green LW, Kreuter MW. Health promotion planning : an educational and environmental approach. Mountain View, CA: Mayfield Pub. Co.; 1991.

[19] Weir C, McLeskey N, Brunker C, Brooks D, Supiano MA. The role of information technology in translating educational interventions into practice: An analysis using the PRECEDE/PROCEED model. J Am Med Inform Assoc. 2011;18(6):827-34.

[20] Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *Am J Public Health*. 1999;89(9):1322-7.

[21] Etingen B, Patrianakos J, Wirth M, et al. TeleWound practice within the Veterans Health Administration: Protocol for a mixed methods program evaluation. *JMIR Res Protoc.* 2020;9(7):e20139.

[22] Donabedian A. The quality of care: How can it be assessed? *JAMA.* 1988;260(12):1743-8.

[23] Watkins KE, Keyser DJ, Smith B, et al. Transforming mental healthcare in the Veterans Health Administration: A model for measuring performance to improve access, quality, and outcomes. *J Healthc Qual.* 2010;32(6):33-43.

[24] Pawson, R, Tilley, N. *Realistic evaluation*. Thousand Oaks, CA: Sage Publications, 1997.

[25] Hulen E, Edwards ST, Poppe AP, Singh MK, Shunk R, Tuepker A. Creating change, challenging structure: Graduate and faculty perspectives on the implementation of an interprofessional education program in veterans affairs primary care. *J Interprof Care.* 2020;34(6):756-62.

Standards of High-Quality Evaluation

Standards and guidelines can be useful in helping evaluators navigate ethical issues and determining best practices through the course of the evaluation. To assess the quality of evaluation activities, the Office of Management and Budget describes several characteristics of high-quality evaluations[26] to help agencies meet the requirements of the Evidence Act:

Standard	Concept	Key Questions
Relevance and Utility	Evaluation should serve information needs of intended users. Results should be easy to interpret and presented in a timely fashion to allow for actionable decisions based on needs of stakeholders	*Are stakeholders' questions effectively addressed by the evaluation?*
Rigor	Evaluations should be conducted using best-supported evaluation practices (e.g., inclusion of a comparison) and conducted by experienced evaluators with relevant skills and education. Caveats that may impact evaluation results interpretation should be clearly stated and explained.	*Will the evaluation be conducted using the best-available personnel and evaluation methods?*
Independence and Objectivity	Evaluators should conduct the evaluation in a manner free from influences that could affect their objectivity, impartiality, or professional judgement	*Will the evaluation be conducted in a manner free from bias or outside influence?*
Transparency	The process of planning, carrying out, and reporting of evaluation results should be clear, well-documented, and obvious to outside observers. Planned work, including dissemination and application, should be clearly documented prior to beginning evaluation activities.	*Is the evaluation plan clear and easy to understand for persons not involved in the evaluation process?*
Ethics	Evaluation staff and stakeholders should consider the rights, safety, and privacy of participants and carry out the evaluation in an ethical manner consistent with current professional standards of equity, fairness, and justice.	*Will the evaluation be carried out in a manner that considers participant impact and professional standards?*

Should I Use Internal or External Evaluators?

The Evidence Act mandates that evaluators are trained to "properly plan, implement, manage, and/or oversee evaluation activities and evaluations." Multiple options for acquiring evaluation resources include interagency agreements, federal awards, or

---

[26] Office of Management and Budget. Phase 4 Implenmention of the Foundation for Evidence-Based Policymaking At of 2018: Program Evaluation Standards and Practices. Memorandum M-20-12. March 10, 2020. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf

other agreements to support planning and implementation, with the key being the use of "qualified evaluators."

With that in mind, an additional consideration involves determining if an internal or external evaluator is most appropriate. Internal evaluators (i.e., persons already affiliated with the program) may have preexisting relationships with stakeholders that allow for greater cooperation, intimate knowledge of the program of interest, access to relevant data, and lower costs associated with carrying out the evaluation; however, internal evaluators may be less objective, more easily experience competing demands and have less formal training in program evaluation. External evaluators (persons without an affiliation with the program) may be more objective, bring new perspectives, have more specialized training in program evaluation, and more dedicated time. At the same time, they may also be more costly, unfamiliar with the setting, and have a more challenging time obtaining staff buy-in.

Designing and Carrying Out a Program Evaluation

An evaluation needs to be informed by local context, shaped to meet the needs of stakeholders, employ rigorous methods, and leverage the best available data. Regardless of the program, framework, or evaluation method, program evaluations should include the following seven steps, which have been adapted from recommendations put forth by the Centers for Disease Control and Prevention (CDC).[27]

> **(1) Engage stakeholders**
>
> **(2) Describe the program and problem**
>
> **(3) Develop an evaluation plan**
>
> **(4) Gather the best available evidence to assess evaluation questions**
>
> **(5) Conduct analyses and form conclusions**
>
> **(6) Discuss findings and recommendations with stakeholders**
>
> **(7) Assess changes from evaluation findings (on treatment programs)**

Despite being numbered sequentially; the process is not always linear. Each step informs and is informed by the other steps in the process.

Step 1: Engage stakeholders

The first step in designing a program evaluation involves seeking out all relevant persons who are involved with the program or who could be affected by the results of the program evaluation. Stakeholders generally include Veterans, providers and other frontline staff, administrators within the program, VA program office leadership, community members, and even family members of health care consumers. Connecting with stakeholders allows evaluators to gather information that can inform the evaluation approach, including measure selection, variable operationalization, intervention conceptualization, and framing of results. The goal of this step is to understand

---

[27] Centers for Disease Control and Prevention. Evaluation Steps. April 9, 2021. Retrieved from
https://www.cdc.gov/eval/steps/index.htm

contextual factors that may impact the program, obtain feedback on measures that best capture the outcomes of interest, and understand priorities to help inform the creation of evaluation products that support uptake and application.

Step 2: Describe the program and problem

Effective evaluation requires an accurate picture of how a program is currently functioning or intended to function as well as factors that impact the outcomes of interest. A useful tool for ensuring an accurate understanding of program functioning is a conceptual, or **Logic Model**. This model is directly informed by program information gathered from stakeholders and includes specific programmatic factors and essential components: goals of the program, the mechanism through which the program is understood to affect positive change, resources available to the program, contextual factors impacting program functioning, and the current day-to-day activities of the program. It is also important to consider information from available research relevant to the evaluation question, setting, and key consumer groups as another source of information to inform the description of the program and problem. These models can help highlight content areas in need of further understanding by the evaluators as well as elucidate opportunities for data collection and interpretation. The CDC provides useful information[28] for the development of logic models.

Step 3: Develop an evaluation plan

Evaluations need to be structured to assess the areas of greatest priority. As such, evaluation plans should incorporate the main questions of interest, important contextual factors, evaluation goals, and resource limitations that could impact the evaluation process or interpretation. These plans should be shaped by the timeline for execution of the evaluation, the goals of potential users for the resulting information, and the expected ways evaluation results could be used to address the problem of focus.

An evaluation plan is the roadmap for the evaluation. It should be detailed but realistic about what can be accomplished within the timeframe and with the resources available. Careful planning early on will reduce the risk of missed opportunities to collect key information and ensure barriers/challenges have been anticipated.

An evaluation plan should include a description, steps, or protocols for each of following:

(1) **Evaluation Question(s):** Determine where to focus the evaluation—clarify the main purpose of the evaluation. The purpose will be informed by stakeholder priorities and/or identified through a needs assessment.

(2) **Program Elements to Monitor:** Describe/identify the program's core components or key actions that will be used to inform answers to key questions and define the pathways from the component/activity to the outcome.

---

[28] Centers for Disease Control and Prevention. Evaluation Guide: Developing and Using a Logic Model. Retrieved from https://www.cdc.gov/dhdsp/docs/logic_model.pdf

(3) Relevant Outcomes: Identified outcomes should be measurable and linked to the elements in #2. Specify the relationship between the core elements and your targeted outcomes.

(4) Appropriate study population and comparison groups: Identify the population that is most relevant to the problem and would potentially benefit most from the evaluation study. Identify a comparison group, ideally via randomization of individuals, or more practically, groups of individuals by site, region, etc. When randomization is not feasible or advisable, a non-randomized comparison group is the next best option, especially if the comparison group has similar characteristics to the group receiving the program. For a practical summary of study designs see the QUERI Implementation Roadmap.[29]

(5) Data Collection and Analysis Plan: Identify the data sources from which you will "collect" your outcomes and other variables. Determine how you will access the necessary data and how the data will be processed and analyzed. Make sure to understand the definitions used by the source and determine if the definition matches closely enough to the definitions of your relevant outcomes. Consider data sources such as policies and directives, observations of the process, feedback from consumers or staff in addition to more traditional data in the form of numerical, systematically collected information in a database. VA REDCap is a secure web platform for building and managing online databases and surveys. NOTE: Primary data collection requires a great deal of effort – determine if quality data already exist and are accessible BEFORE planning to collect surveys or other data yourself.

(6) Resources: Resource availability may impact your ability to evaluate some aspects of your questions of interest. Be creative and determine what evaluation questions you can address given the resource constraints you face. There is still a considerable benefit to incrementally evaluating a program.

(7) Dissemination Plan: Identify who will receive evaluation findings and how frequently information should be shared. Include details such as the mode of information sharing (e.g., formal presentation/report, one-page fact sheet, manuscript).

(8) Sustainability Plan: Identify resources needed to sustain the program, including those for ongoing support and training. Describe the approach you will take to integrate the program into standard care. Continued stakeholder buy-in and explicit support cannot be emphasized enough—plan for how this will be maintained to increase the likelihood that a program will be sustained.

---

[29] Goodrich DE, Miake-Lye I, Braganza MZ, Wawrin N, Kilbourne AM. Quality Enhancement Research Initiative. QUERI Roadmap for Implementation and Quality Improvement. United States Department of Veterans Affairs Veterans Health Administration Office of Research and Development Health Services Research and Development, 2020. Retrieved from https://www.queri.research.va.gov/tools/roadmap/

Keep in mind that evaluations are not set in stone. Plan in as much detail as possible early on but remain flexible as unexpected challenges can arise (e.g., staffing changes, policy/directive changes).

Step 4: Gather best available evidence to assess evaluation questions

To effectively use data to assess program functioning, evaluators need to match program conceptualization to available data, keeping in mind the quality, location, and limitations of each dataset and variable. As you are developing your evaluation plan, you will have operationalized key program factors, linked each to relevant data sources, and identified any data quality concerns that might impact the evaluation design and results interpretation. Lastly, plan so that the process for obtaining access permissions for each dataset of interest, as well as storing and data management requirements are in compliance with relevant data privacy and security standards.

Step 5: Conduct analyses and form conclusions

Once relevant program data have been collected, evaluators need to use appropriate techniques to conduct analyses, interpret findings, and formulate recommendations. Analyses will differ depending on the evaluation design and the type of data collected. For example, qualitative evaluations utilizing interview and/or focus group data collection methods may employ thematic analysis.[30] Interpretation of results should take into account program factors (e.g., priorities, treatment approach, programmatic environment), evaluation limitations (e.g., focus on one area over others, evaluation scope and capacity), data limitations (e.g., factors such as the limits of measures included in the evaluation and missing/unavailable data), and other available information sources relevant to the topic (e.g., research articles, other evaluation reports). If there is a comparison between groups receiving versus not receiving the program, then additional sociodemographic data should be ascertained to ensure that program differences are distinct from differences in the subpopulations receiving or not receiving the program. All recommendations should be directly linked to findings and should include any caveats based on limitations of approach to avoid misinterpretations of findings.

Step 6: Discuss findings and recommendations with stakeholders

It is important for evaluators to share conclusions based on evaluation findings with stakeholders to support any needed program changes. It can also be helpful to share interim results with stakeholders as a way to keep them engaged in the evaluation process. When presenting findings, evaluators should be clear in describing the evaluation and interpretation process to allow stakeholders to draw their own conclusions. A common outcome of this step is the creation of a summary evaluation report that presents findings, interpretations, and recommendations intended to inform decisions about program functioning. Tables and figures are useful in making both quantitative and qualitative results easier to interpret. It may also be useful to create multiple products (e.g., reports, manuscripts, slides, etc.) to help share evaluation findings and recommendations to a range of stakeholder groups.

---

[30] Braun, V, Clarke, V. Using Thematic Analysis in Psychology. *Qual Res Psychol*. 2006;3(2):77–101.

Step 7: Assess changes from evaluations findings (on treatment programs)

Program evaluations often include steps to continue to monitor program functioning after initial evaluation efforts are completed. Ongoing monitoring of program functioning is especially valuable if changes were made to the program based on recommendations from the initial evaluation. While ongoing monitoring is not a standard part of all evaluations, it is strongly recommended that evaluation teams take steps to support stakeholders and transition ownership of ongoing monitoring by providing training to program staff on how to continue program monitoring efforts independently. Evaluators can also build in sustainability of program monitoring as an expectation for next steps while the primary evaluation is wrapping up. For example, the evaluator might decrease the intensity of the evaluation approach or shift the focus of the evaluation to assess the which recommendations from the initial evaluation are effectively implemented into practice. Plans for ongoing evaluation should be included in all summary products, including recommendations for how monitoring is to be conducted and what changes are being made to program functioning.

Choosing a Methodological Approach that Fits your Evaluation

When designing an evaluation, it is important to ensure your evaluation design matches the main questions of interest. One key consideration is the extent to which the main evaluation questions are better answered via the use of qualitative or quantitative approaches. A brief distinction between these two approaches is summarized below:[31]

Characteristic	Quantitative Methods	Qualitative Methods
Data description	Quantitative data is data that can be counted, measured, or otherwise expressed in numbers	Qualitative data is descriptive, concerning observable phenomena from the informants' perspectives
Data collection methods	Commonly collected through closed-ended surveys or questionnaires, pre-tests and post-tests, observation, or by gathering clinical numeric data	Commonly collected from focus groups, open-ended interviews, questionnaire responses, first-hand observation notes, documents, or artifacts
Analysis	Analysis involves the use of statistical methods to describe, summarize, and compare data; statistical analyses can range from basic descriptive statistics to complex inferential analyses	Analysis involves identification of themes, coding, clustering similar data, and/or reducing data to meaningful and important patterns
Central evaluation questions	Focuses on "what" and "how many"	Focuses on "why" and "how"
Data presentation	Tells a story with numbers (numerical)	Tells a story with words (narrative)
Best application	Used to test hypotheses, assess cause and effect, and make predictions	Used to develop hypotheses, elicit "insider" or stakeholder perspectives, and/or understand and interpret behavior

---

[31] Centers for Disease Control and Prevention. Evaluation Steps. April 9, 2021. Retrieved from https://www.cdc.gov/eval/steps/index.htm

Mixed Methods

Mixed methods designs are becoming more commonly applied in program evaluation. Mixed methods are characterized by an integration of quantitative and qualitative approaches to collecting and analyzing information. This integration permits a more complete examination and assessment of the impact of a program that would be possible with a single approach.[32]

In deciding whether a mixed methods approach is appropriate for your evaluation, consider the following examples of situations where a mixed methods approach is needed:

(1) Currently available data is not sufficient to address every evaluation question

(2) A need exists to explain a set of initial results

(3) There is a desire to generalize exploratory findings

(4) A need exists to enhance an evaluation with an embedded second method

(5) A need exists to understand the program objectives through multiple evaluation phases

The Value of (Careful) Creativity in Evaluation Design

Careful creativity when designing an evaluation can be useful, especially when there are limitations to available data and evaluation techniques. When you have flexibility in defining your metrics of interest, consider creating outcome proxies that allow you to assess an area of interest that may not be directly captured by available data. For example, available data may lack measures of Veteran symptom levels but include counts of inpatient care days. By shifting focus to condition severity rather than symptom levels, an evaluation could assess whether a treatment reduced condition severity (i.e., inpatient care days), even in the absence of available symptom-level information.

Creativity and flexibility in evaluation design can also be helpful when deciding on how to operationalize measures and how evaluation results could be used to address questions that may be beyond the initial focus of the evaluation. For example, a clinical intervention designed to provide primary care via telehealth could have a primary outcome interest of customer safety and access to care—but this could also impact provider job satisfaction and productivity. Although the initial request for a program evaluation could focus on consumer outcomes, a slightly broader view when designing the evaluation could provide a more comprehensive view of the multiple ways that a change in clinical practice impacts the health care system and process.

It is important, though, to be sure that results are not interpreted to imply causal relationships that extend beyond the limitations of the data and evaluation approach, or that proxies are over-interpreted in terms of what they reflect. For example, increased treatment participation in Veterans with elevated levels of disability may reflect appropriate investment in treatment for Veterans with more serious conditions. It could

---

[32] Creswell JW, Clark VL. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage Publications; 2018.

also suggest an increased attention to screening, diagnosis, and follow-up by providers for Veterans who are seen more often (or both). A focus on association over causation will help reduce misinterpretation of results when using measure proxies and flexible definitions of measures of interest.

Using Existing Clinical and Administrative Data

Both clinical and administrative data can be used to monitor outcomes during an implementation or quality improvement initiative. Clinical data may provide a more complete view of a patient's health and health care. Administrative data is frequently collected for program and health care operations management and can also provide additional information for evaluation. However, using clinical and administrative data to support program evaluations can be complicated, depending on the source and structure of the data, as this data was collected for purposes unrelated to your evaluation. Users of existing data should be cognizant of the purposes for which it was originally collected.

VA-based evaluation plans often rely heavily on available health and health care data collected in the process of providing care to Veteran patients. Consult the VHA Data Portal for a list of centrally available data. Some common VA data sources are:

- Assistant Deputy Under Secretary for Health (ADUSH) Enrollment Files
- Cardiac Assessment Reporting and Tracking System (CART)
- Centers for Medicare & Medicaid Services (CMS) claims data
- CMS Data for Veterans: Medicare, Medicaid, Patient Assessment, Healthcare Effectiveness and Information Set (HEDIS)
- External Peer Review Program (EPRP), Analytics and Performance Improvement
- Home and Community Based Quality Care Initiative
- Inpatient Evaluation Center (IPEC)
- Managerial Cost Accounting (MCA) National Data Extracts (NDEs)
- Mental Health Program Evaluation Centers (e.g., VA Serious Mental Illness Treatment, Resource, and Evaluation Center (SMITREC))
- Mortality Data Repository (MDR)
- MyHealtheVet usage data
- National Center for Homeless Veterans Registry
- Pharmacy Benefits Management Database (PBM)
- Planning Systems Support Group (PSSG) Geocoded Enrollee Files
- Primary Care Management Module (PCMM) Web
- Survey of Health Care Experiences of Patients (SHEP)
- United States Renal Data System (USRDS)

- United States Veterans Eligibility Trends and Statistics (USVETS)

- VA Airborne Hazards and Open Burn Pit Registry

- VA All Employee Survey (AES), Voice of VA Survey

- VA Central Cancer Registry (VACCR)

- VA Cerner Millennium

- VA Surgical Quality Improvement Program (VASQIP)

- Veteran Administration Site Tracking (VAST) Database

- VHA Corporate Data Warehouse (CDW)

- VHA Quality Improvement Initiative (VQuIP)

- VHA Vital Status File (VSF)

Evaluators are strongly encouraged to connect with VA regional leaders (e.g., Veterans Integrated Service Network (VISN) or facility leaders) to assess the types of data and metrics that they are accountable for, in order to inform the creation of an evaluation plan. The Health Services Research & Development's (HSR&D's) VA Information Resource Center (VIREC) is also available to provide consultation on data sources, structure, and availability.

Consider the below recommendations before utilizing administrative data for evaluation.

Figure 2: Dos and Don'ts when using existing clinical and administrative data

DO
Consider multiple data sources—the desired data may not be in one place
Ask stakeholders for feedback about available data relevant to the evaluation
Shape your evaluation and outcome definitions around available data
Work to gain familiarity with data sources before settling on final measure definitions
Consider varied data sources that may conceptualize outcomes in different ways
Include data limitations in formation of conclusions and recommendations
Expect variability in data completeness across consumers, measures, and time period
If you have questions about available VA data sources, contact VIReC HelpDesk

DON'T
Extend findings beyond limitations of data
Assume findings will generalize
Decide information is unavailable if not found in first (or second) data source
Remain rigid in your definition of concepts if information to support definition is unavailable
Develop effortful data collection plans if administrative data not checked first

Data Analysis

Data analysis is "the process of organizing and classifying the information you have collected, tabulating it, analyzing it, comparing the results with other appropriate information, and presenting the results in an easily understandable manner."[33] When planning and ultimately carrying out the analysis of the evaluation data and interpretation of quantitative results, evaluators should consider the role of both statistical and substantive significance of findings and design their evaluation approach to assess both areas.

Statistical Significance concerns null hypothesis significance testing and a focus on group differences. Statistical significance, quantified by the familiar "*p-value,*" is the probability that your results are due to chance rather than real program effect. *P-values* are dependent on sample size (i.e., the number of observations/patient visits). Given very large numbers of Veterans in VA, a very small effect can be statistically significant but not practically significant. For example, an increase of 0.2 encounters could be statistically significant; however, programs do not offer partial encounters, making this finding less useful for informing care delivery.

Substantive Significance focuses on "the degree to which interested parties (scientists, practitioners, managers, consumers, decision makers, the public at large, etc.) would consider a finding important and worthy of attention and possibly action."[34] Substantive significance can be determined based on other evaluations in the literature reporting the impact of a similar program, but it is more often that clinical judgement will decide if a finding is substantively significant.[35] There are several measures of effect size that can be used to estimate substantive significance, including Cohen's *d*.[36] Evaluators should consider their specific context when deciding to apply standard conventions regarding effect size interpretation. In some cases, small effect sizes can have important implications.

Collection and analysis of qualitative data should also be conducted in a manner consistent with best-practice recommendations and generally should focus on the use of well-supported approaches, such as content analysis, narrative analysis, and the Constant Comparative Method. More information related to qualitative data methods and analytic approaches is available via CDC's Field Epidemiology Manual.[37]

Data Interpretation Tips

Interpretation of evaluation results should always be framed by the evaluation questions themselves. Important considerations when interpreting results include:

---

[33] MacDonald G, Starr G, Schooley M, Yee SL, Klimowski K. Introduction to program evaluation for comprehensive tobacco control programs. Atlanta (GA): Centers for Disease Control and Prevention; 2001.

[34] Kelley K, Preacher KJ. On effect size. *Psychol Methods.* 2012;17(2):137-52.

[35] Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: Clinical versus statistical significance. *Perspect Clin Res*. 2015;6(3):169-70.

[36] Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic; 1988.

[37] Wolff B, Mahoney F, Lohiniva AL, Corkum M. Chapter 10. Collecting and Analyzing Qualitative Data. In Rasmussen SA, Goodman RA. *The CDC Epidemiology Manual.* New York: Oxford University Press, 2019. Available at: https://www.cdc.gov/eis/field-epi-manual/chapters/Qualitative-Data.html.

- **Stakeholder audience:** What questions do they want answered?

- **Limitations of the evaluation design and data:** What biases were introduced? How do these factors affect the validity and reliability of the results?

- **Alternative explanations for your findings:** Could unmeasured factors have contributed to your findings?

- **How do these results compare to evaluations of similar programs using similar methods?**

- Have **different evaluation approaches focusing on similar questions found similar results?**

- How do **results compare to theories** supported by other evaluations?

- Are **results different from what was expected** by stakeholders/evaluators? If so, what are some factors that may have led to this difference?

A consideration of these factors will help to ensure that evaluation interpretations are comprehensive, balanced, and more easily applied in ways that address the questions which motived the evaluation.

Partnered Evaluations

There is a growing movement toward the development of partnered evaluation designs in which two (or more) organizations collaborate in the development of an evaluation that is codesigned and executed with the goal of answering questions related to mutual high-priority areas. Such partnerships can support the creation of high-quality evaluations through sharing of funding, data, and staff resources, which allows for a larger and more rapid evaluation than might have been possible if undertaken by one organization alone. In VA, QUERI supports national and regional partnered evaluations, co-funded by QUERI and a VA operations partner (e.g., National Program Office, VISN leader), through the QUERI Partnered Evaluation Initiative (PEI) mechanism.

In 2017, QUERI conducted a review of its existing PEI evaluations. A four-step process emerged from this review that summarizes best practices related to the development of effective operational/investigator partnered evaluation. Many of these steps overlap with steps previously presented to support the development of effective program evaluations:

Step 1: Create evaluation project plan and team

1a: Develop evaluation questions and planned products based on areas of VA priority

The initial step in the process of conducting a QUERI-supported program evaluation project is the creation of clear evaluation questions that address an area of concern or problem within VA. Potential evaluation team members are encouraged to familiarize themselves with current areas of national and regional priority to aid in the development of focused, applied evaluation plans that result in actionable products.

1b: Identify potential collaborators

In some cases, it is an easy process to identify potential collaborators for an evaluation project based on previous work experience, shared professional network, or reputation. In other cases, it may be more challenging to identify persons with relevant expertise and interest to help move a planned evaluation project forward. A consideration of high-level operational stakeholders (e.g., organizational leadership) and persons with specialized knowledge of the main evaluation area (e.g., research investigators) can be helpful when identifying potential collaborators to participate in the evaluation team.

1c: Develop a project plan based on VA evaluation best practices

The information summarized in this guide represents best practice recommendations for the development of an evaluation consistent with VA expectations. Persons interested in conducting a program evaluation in VA are encouraged to review this guide and use its information to help with the development of an evaluation project plan consistent with these recommendations.

1d: Review available data sources relevant to the evaluation topic and develop Data Use Agreement with partner

Information related to VHA administrative data resources is presented previously in this guide ("17"). Health system operations leaders and research investigators often have additional internal datasets or preferred data resources for use when evaluating their programs. Because of this, partners are encouraged to familiarize themselves with previous work that has been done by their potential partner as well as discuss potential data sources with their partner for inclusion in their evaluation plan.

Data Use Agreement: If you plan to use any data from an operations partner, a Data Use Agreement is required, which outlines appropriate uses of the data as well as process of reviewing products based on the data for publications, presentations, etc.

1e: Develop work agreements with collaborators

Partnered collaboration can be challenging, given differences in perspective, priority, and approach. To understand some ways that evaluation partners overcome these differences, the QUERI Center for Evaluation and Implementation Resources (CEIR), conducted interviews of operations partners and QUERI investigators collaborating on QUERI Partnered Evaluation Initiatives. The interviews sought to identify barriers and facilitators of effective collaboration between research and operations partners, establish best practices to bridge differences and maximize the value of multiple perspectives, and develop a set of recommendations for future partnered evaluation work. The following page summarizes lessons learned related to the characteristics supportive of effective working relationships and best practices designed to support these characteristics:

Characteristic	Definition	Best Practice Supportive of Adaptive Functioning
Trust	Mutual appreciation for the skills and judgement of partners	Explain thought process; reduce use of jargon; develop new partnership among persons who have previously collaborated; use a clearly defined, mutually agreed upon work plan
Shared Understanding	Appreciation for differing perspectives and priorities of partners; agreed-upon approach to conducting the project work	Open communication of preferences and pressures; formal drafting of project plan
Collaboration	Inclusivity; ongoing discussions related to project challenges; open to negotiation; unified team mentality	Regular plans for communication; multiple points of contact to support project teams; in-person meetings; regular meetings to review project performance
Adaptability	Able to shift approach based on changes in priorities	Planning for flexibility in later periods of project plan; regular review of project performance and areas of concern; regular discussion of outside pressures that might impact project
Mutual Benefit	Products and processes designed to meet the differing needs of different partners	Regular review of planned products; clear plan for a variety of products differentially designed around partner priorities; focus on direct translation of projects to clinical care; development of parallel products from same project data
Products	Presence of clear plans for product dissemination, product ownership, and partner roles related to creating and disseminating products	Creation of clear plans for product development, ownership, and dissemination; clarity about the primary objectives of the partnership help to reduce disagreement about product development or prioritization
Investment	Mutual, active participation between partners in supportive ways that demonstrated active investment in the partnership	Open, ongoing discussions related to funding and other needed resources; discussion of collaboration opportunities beyond scope of project; regular participation in project meetings and timely response to project communications
Leadership Support	Active, informed, supportive involvement of partner supervisors	Include leadership team in initial discussions related to project creation; provide regular updates on project performance; include project performance in performance evaluation

Step 2: Formalize evaluation project plan

2a: Determine whether your evaluation meets the definition of non-research

Organizations can collaborate in multiple ways that involve different methods for data collection, analysis, and planned product creation. One important consideration is whether the planned evaluation will be designed in a way that is focused solely on quality improvement in VA or for generalizability. Guidance related to considerations that can shape this decision in VA are included previously in this guide ("Is my evaluation considered research or non-research?") and are based on definitions from

ORD Program Guide 1200.21.[38] For evaluations that are designed in collaboration with a VA operations partner for quality improvement purposes, it is important to document non-research status.

VA-Specific Guidance
The process for documenting non-research status in VHA is described in the following: Jeans CK, Kilbourne AM, Booker F, Braganza MZ. VA Quality Enhancement Research Initiative (QUERI): Everything you need to know about a QUERI non-research ("QI") protocols. 2019. https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/3669-notes.pdf

2b: Draft a Memorandum of Understanding (MOU) between all persons involved in the evaluation

Within evaluation partnerships, there are many opportunities for misunderstandings related to planned evaluation activities, roles, timelines, and planned products. One effective tool to help reduce such misunderstandings and ensure that there is a clearly agreed-upon plan for the collaboration is a **Memorandum of Understanding (MOU)**. An MOU is a formalized agreement between all members of the evaluation team (and any relevant outside groups, such as funding bodies who may be providing monies to support the evaluation) to ensure alignment on key milestones and products, funding and resources, and the project timeline**.** The MOU should state the scope of the work to be done, the funding commitment by the operations partner and any funding bodies such as VA HSR&D/QUERI, and the timeline for completion of project deliverables. In the case of evaluation partnerships that take place outside of the QUERI Partnered Evaluation Initiative mechanism, the MOU can be modified to fit the characteristics of evaluation partners and planned work.

Step 3: Conduct the evaluation in alignment with evaluation best practices and VA policies

Make sure to notify the facility leadership, union representatives (process described here), and the Associate Chief of Staff for Research at evaluation sites before conducting any surveys or interviews. Both research and non-research evaluations might involve several VA facilities where data collection (e.g., organizational surveys) is done centrally at a single facility.

Step 4: Disseminate evaluation results

Often, evaluations present opportunities for the creation of manuscripts and conference presentations as mechanisms through which to disseminate results. In these situations, journals will often ask for documentation that the work underwent review, and in this case, it would be appropriate to submit the non-research documentation letter. Within

---

[38] Department of Veterans Affairs Office of Research and Development. Program Guide 1200.21. VHA Operations Activities that May Constitute Research. January 9, 2019. Retrieved from https://www.research.va.gov/resources/policies/ProgramGuide-1200-21-VHA-Operations-Activities.pdf

partnered evaluation projects, it can be helpful to have clear front-end discussions of planned products and dissemination plans to avoid confusion about how to best share lessons learned. Such discussions are also important because partners may have differences in perspective of which products will provide the most organizational value to them based on their performance expectations.

How do I find funding for my evaluation?

In VA, there are two principal sources of funding for program evaluation—HSR&D and QUERI—both of which accept program evaluation applications that undergo scientific peer review. VA employees have the opportunity to apply for HSR&D/QUERI funding to support evaluation of programs that affect Veterans. HSR&D funding mechanisms emphasize innovation and support research that takes a long look at a problem or challenge in VA. Completely funded by HSR&D, these evaluations require Institutional Review Board review. The focus of research-funded evaluations includes questions that seek to inform initiatives affecting Veterans as well as expand the knowledge base of health services research. In contrast, QUERI funding supports rigorous evaluations of programs or policies selected and co-funded by an operations partner and address VA national priorities to ultimately improve Veteran care. QUERI supports rigorous evaluations that are aligned with time-sensitive VA priorities and meet the definition of non-research, as defined by Program Guide 1200.21. Most QUERI evaluations are primarily directed and funded by VA operations partners (e.g., VA National Program Office, VISN).

The QUERI Partnered Evaluation (PEI) request for applications (RFA), part of the QUERI Global Merit Review RFA, is one of the principal sources of methods and evaluation guidance for VA's response to the Evidence Act through the Office of Management and Budget. Examples of current QUERI Partnered Evaluations and operations partners can be found on the QUERI website.

There are key advantages for VA operations partners (e.g., Program Offices, VISNs) in collaborating with HSR&D/QUERI on evaluations:

- The evaluation plan is subject to a rigorous scientific review process (thus enabling relevant and state-of-the-art evaluation methods) and is nationally recognized as peer-reviewed scholarship through the National Institutes of Health (NIH) Reporter.

- For partnered evaluations, operations partners that commit additional funds to the evaluation work can add these to their budget, hence showing commitment to evidence-based policy as mandated by the Evidence Act and Office of Management and Budget.[39]

- Results from the evaluation can inform what resources, policies, or infrastructure are required to sustain the program or policy, if demonstrated effective.

For more details on how to apply for funding, please review the VA HSR&D website.

---

[39] Bridgeland J, Orszag P. Can government play moneyball? How a new era of fiscal scarcity could make Washington work better. *The Atlantic.* 2013.

Bibliography

Atkins D, Kilbourne AM, Shulkin D. Moving from discovery to system-wide change: The role of research in a Learning Health Care System: Experiences from three decades of health systems research in the Veterans Health Administration. *Annu Rev Public Health*.2017;38:467–87.

Braun, V, Clarke, V. Using Thematic Analysis in Psychology. *Qual Res Psychol*. 2006;3(2):77–101.

Bridgeland J, Orszag P. Can government play moneyball? How a new era of fiscal scarcity could make Washington work better. *The Atlantic.*2013. Available at: https://www.theatlantic.com/magazine/archive/2013/07/can-government-play-moneyball/309389/. Accessed March 15, 2021.

Centers for Disease Control and Prevention. Evaluation Guide: Developing and Using a Logic Model. Available at: https://www.cdc.gov/dhdsp/docs/logic_model.pdf. Accessed March 15, 2021.

Centers for Disease Control and Prevention. Evaluation planning: What is it and how do you do it? Available at: https://www.cdc.gov/healthcommunication/pdf/evaluationplanning.pdf. Accessed March 15, 2021.

Centers for Disease Control and Prevention. Evaluation Standards. Available at: https://www.cdc.gov/eval/standards/index.htm. Reviewed April 9, 2021. Accessed March 15, 2021.

Centers for Disease Control and Prevention. Evaluation Steps. April 9, 2021. Available at: https://www.cdc.gov/eval/steps/index.htm. Accessed March 15, 2021.

Centers for Disease Control and Prevention. Framework for program evaluation in public health. MMWR 1999;48(No. RR-11).

Clinical and Translational Science Awards Consortium Community Engagement Key Fucntion Committee Task Force on the Principles of Community Engagement. Principles of Community Engagement. Second Edition. June 2011. NIH publication 11-7782.

Capwell EM, Butterfoss F, Francisco VT. Why Evaluate? *Health Promot Pract.* 2000;1(1):15-20.

Cohen AN, Glynn SM, Hamilton AB, Young AS. Implementation of a family intervention for individuals with schizophrenia. *J Gen Intern Med.* 2010;25(1):32-7.

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic; 1988.

Community Matters. (2008). Brief introduction to realist evaluation. Available at: http://www.communitymatters.com.au/gpage1.html

Creswell JW, Clark VL. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage Publications; 2018.

Department of Veterans Affairs Office of Research and Development. Program Guide 1200.21. *VHA Operations Activities that May Constitute Research*. January 9, 2019.

Donabedian A. The quality of care: How can it be assessed? *JAMA*. 1988;260(12):1743-8.

Ellison ML, Mueller L, Smelson D, Corrigan PW, Torres Stone RA, Bokhour BG, Najavits LM, Vessella JM, Drebing C. Supporting the education goals of post-9/11 veterans with self-reported PTSD symptoms: A needs assessment. *Psychiatr Rehab J*. 2012;35(3):209-17.

Etingen B, Patrianakos J, Wirth M, et al. TeleWound practice within the Veterans Health Administration: Protocol for a mixed methods program evaluation. *JMIR Res Protoc*. 2020;9(7):e20139.

Foundations for Evidence-Based Policymaking Act of 2018. Public Law 115-435. 132 STAT 5529 (2018).

Freeman RC, Sukuan N, Tota NM, Bell SM, Harris AG, Wang HL. Promoting spiritual healing by stress reduction through meditation for employees at a Veterans hospital: A CDC framework–based program evaluation. *Workplace Health Saf*. 2020;68(4):161-70.

Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *Am J Public Health*. 1999;89(9):1322-7.

Goodrich DE, Miake-Lye I, Braganza MZ, Wawrin N, Kilbourne AM. Quality Enhancement Research Initiative. QUERI Roadmap for Implementation and Quality Improvement. United States Department of Veterans Affairs Veterans Health Administration Office of Research and Development Health Services Research and Development, 2020. Available at: https://www.queri.research.va.gov/tools/roadmap/

Green LW. Toward Cost-Benefit Evaluations of Health Education: Some Concepts, Methods, and Examples. *Health Educ Behav*. 1974;2(1_suppl):34-64.

Green LW, Kreuter MW. *Health promotion planning : an educational and environmental approach*. Mountain View, CA: Mayfield Pub. Co.; 1991.

Guise JM, Savitz LA, Friedman CP. Mind the gap: putting evidence into practice in the era of learning health systems. *J Gen Intern Med*. 2018; 33:2237-9.

Horwitz LI, Kuznetsova M, Jones SA. Creating a learning health system through rapid-cycle, randomized testing. *N Engl J Med*. 2019;381(12):1175-9.

Hulen E, Edwards ST, Poppe AP, Singh MK, Shunk R, Tuepker A. Creating change, challenging structure: Graduate and faculty perspectives on the implementation of an interprofessional education program in veterans affairs primary care. *J Interprof Care*. 2020;34(6):756-62.

Jeans CK, Kilbourne AM, Booker F, Braganza M. VA Quality Enhancement Research Initiative (QUERI): Everything you need to know about a QUERI non-research ("QI") protocols. 2019. Available at: https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/3669-notes.pdf. Accessed March 15, 2021..

Kelley K, Preacher KJ. On effect size. *Psychol Methods.* 2012;17(2):137-52.

Kilbourne AM, Goodrich DE, Miake-Lye I, Braganza MZ, Bowersox NW. Quality Enhancement Research Initiative Implementation Roadmap: Toward sustainability of evidence-based practices in a learning health system. *Med Care.* 2019;57(10 Suppl 3):S286-93.

Lowe NK, Cook PF. Differentiating the scientific endeavors of research, program evaluation, and quality improvement studies. *J Obstet Gynecol Neonatal Nurs.* 2012;41(1):1-3.

MacDonald G, Starr G, Schooley M, Yee SL, Klimowski K. Introduction to program evaluation for comprehensive tobacco control programs. Atlanta (GA): Centers for Disease Control and Prevention; 2001.

Mattocks KM, Kuzdeba J, Baldor R, Casares J, Lombardini L, Gerber MR. Implementing and evaluating a telephone-based centralized maternity care coordination program for Pregnant Veterans in the Department of Veterans Affairs. *Womens Health Issues.* 2017;27(5):579-85.

Office for Human Research Protections. Revised Common Rule 2018. Federal Register Vol 82 No 12. January 19, 2017.

Office of Management and Budget. Circular No. A-11. Preparation, Submission, and Execution of the Budget. June 2019.

Office of Management and Budget. Phase 4 Implenentation of the Foundation for Evidence-Based Policymaking At of 2018: Program Evaluation Standards and Practices. Memorandum M-20-12. March 10, 2020.

Patton, MQ. *Utilization-Focused Evaluation*, 4th edition. Beverly Hills: Sage Publications; 2008.

Pawson, R, Tilley, N. Realistic evaluation. Thousand Oaks, CA: Sage Publications, 1997.

Office of the Federal Register, National Archives and Records Administration. "45 CFR 46.102 - Definitions.". Government. Office of the Federal Register, National Archives and Records Administration, September 30, 2007.

Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: Clinical versus statistical significance. *Perspect Clin Res.* 2015;6(3):169-70.

Riley WJ, Moran JW, Corso LC, Beitsch LM, Bialek R, Cofsky A. Defining quality improvement in public health. *J Public Health Manag Pract.* 2010;16(1):5-7.

NORC Walsh Center for Rural Health Analysis and University of Minnesota Rural Health Research Center. Rural Health Information Hub. Chapter 5. Evaluation Frameworks. In Rural Care Coordination Toolkit [online]. 2018. Available at https://www.ruralhealthinfo.org/toolkits/care-coordination/5/evaluation-frameworks

Salabarría-Peña Y, Apt BS, Walsh CM. Practical Use of Program Evaluation among Sexually Transmitted Disease (STD) Programs, Atlanta (GA): Centers for Disease Control and Prevention; 2007.

U.S. Department of Health and Human Services Centers for Disease Control and Prevention. Office of the Director, Office of Strategy and Innovation. Introduction to program evaluation for public health programs: A self-study guide. Atlanta, GA: Centers for Disease Control and Prevention, 2011.

Walsh PL, Garbs CA, Goodwin M, Wolff EM. An impact evaluation of a VA geriatric team development program. *Gerontol Geriatr Educ.* 1995;15(3):19-35.

Watkins KE, Keyser DJ, Smith B, et al. Transforming mental healthcare in the Veterans Health Administration: A model for measuring performance to improve access, quality, and outcomes. *J Healthc Qual.* 2010;32(6):33-43.

Weir C, McLeskey N, Brunker C, Brooks D, Supiano MA. The role of information technology in translating educational interventions into practice: An analysis using the PRECEDE/PROCEED model. *J Am Med Inform Assoc.* 2011;18(6):827-34.

Wolff B, Mahoney F, Lohiniva AL, Corkum M. Chapter 10. Collecting and Analyzing Qualitative Data. In Rasmussen SA, Goodman RA. *The CDC Epidemiology Manual.* New York: Oxford University Press, 2019. Available at: https://www.cdc.gov/eis/field-epi-manual/chapters/Qualitative-Data.html. Accessed June 3, 2021.

Additional Evaluation Resources

Evaluation Resources: Methods

- Creswell JW, Clark VL. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage Publications; 2018.
- Pancer SM and Westhaus A. A developmental stage approach to program planning and evaluation. *Eval Rev*. 1989;13:56-77.
- Patton MQ. *Qualitative Research Evaluation Methods*. Thousand Oaks, CA: Sage Publications; 2014.
- QUERI – Quality Enhancement Research Initiative: https://www.queri.research.va.gov/ceir/default.cfm
- Riley WJ, Moran JW, Corso LC, Beitsch LM, Bialek R, Cofsky A. Defining quality improvement in public health. *J Public Health Manag Pract*. 2010;16(1):5–7.
- Rossi PH, Freeman HE, Lipset MW. *Evaluation: A systematic approach*. 6th ed. Thousand Oaks, CA: Sage Publications; 1999.
- Infant and Early Childhood Mental Health Consultation (IECMHC) Toolbox: Research and Evaluation. Available at: https://www.samhsa.gov/iecmhc/toolbox/research-evaluation

- USAID. Technical Note – Conducting Mixed-Methods Evaluations. Version 1. 2013. Available at: https://www.usaid.gov/sites/default/files/documents/1870/Mixed_Methods_Evaluations_Technical_Note.pdf
- CDC Evaluation Methods Resources
  - https://www.cdc.gov/eval/index.htm
  - https://www.cdc.gov/healthcommunication/pdf/evaluationplanning.pdf
  - https://www.cdc.gov/std/Program/pupestd/Selecting%20Data%20Collection%20Methods.pdf
  - MacDonald G, Starr G, Schooley M, Yee SL, Klimowski K, Turner K. Introduction to Program Evaluation for Comprehensive Tobacco Control Programs. Atlanta (GA): Centers for Disease Control and Prevention; 2001.
  - Centers for Disease Control and Prevention. Brief 19: Analyzing Qualitative Data for Evaluation, 2018. Available at: https://www.cdc.gov/healthyyouth/evaluation/pdf/brief19.pdf
  - Centers for Disease Control and Prevention. Brief 20: Analyzing Quantitative Data for Evaluation. Evaluation Brief 20. Analyzing Quantitative Data for Evaluation, 2018. Available at: https://www.cdc.gov/healthyyouth/evaluation/pdf/brief20.pdf

Evaluation Resources: Frameworks

- Fynn JF, Hardeman W, Milton K, Jones AP. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. *BMC Public Health*. 2020;20(1):1000.
- Patton MQ. *Qualitative Research Evaluation Methods*. Thousand Oaks, CA: Sage Publications; 2014.
- Markiewicz A, Patrick I. *Developing Monitoring and Evaluation Frameworks*. Thousand Oaks, CA: Sage Publications; 2016.
- Rossi PH, Freeman HE, Lipset MW. *Evaluation: A systematic approach*. 6th ed. Thousand Oaks, CA: Sage Publications. 1999.
- CDC Evaluation Frameworks
  - MacDonald G, Starr G, Schooley M, Yee SL, Klimowski K, Turner K. Introduction to Program Evaluation for Comprehensive Tobacco Control Programs. Atlanta (GA): Centers for Disease Control and Prevention; 2001.
  - Centers for Disease Control and Prevention. Framework for program evaluation in public health. *MMWR*. 1999;48(No. RR-11).