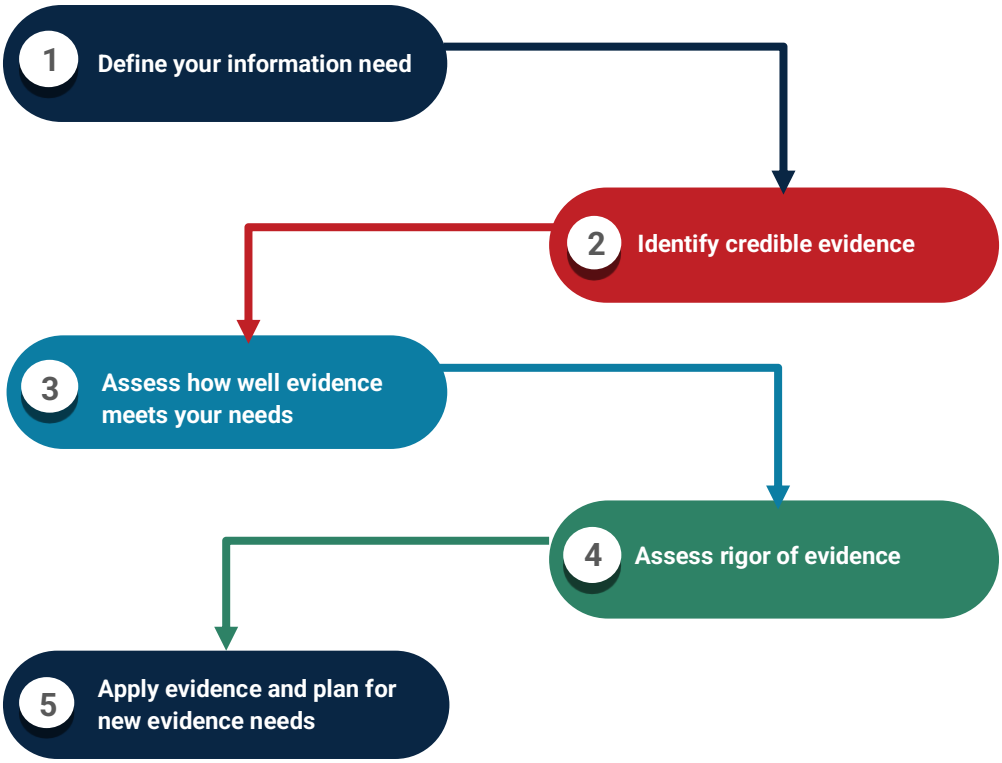
How to Use Available Evidence to Inform Decision-Making

Decision makers need the best possible evidence to use when designing or considering changes to programs, policies, or regulations. Understanding whether such activities are working as intended, for example, can help inform decisions to fund specific services or expand selected pilot programs. However, available evidence often reflects a specific moment, context, or intervention that is similar to, but not exactly the same as, what you are seeking to address. Different sources of evidence may lead to different conclusions, results may not be conclusive, or the data available or methods used may have limitations.

To address these realities, this guide describes the key steps to use and apply available evidence in decision-making, including when that evidence is imperfect.



1 Define your information need.

Be clear about what you need to know and how the credible and relevant evidence you seek will be used for policy or program decisions. Your question or problem statement will guide every step of the process, beginning with identifying the best available evidence, assessing the rigor and applicability of that identified evidence, and making recommendations based on the evidence. Your question or problem statement will inform what you need to know about the program, policy, or regulation; the population and outcomes of interest; and context (e.g., geography).

For example, we might ask:

- ***To what extent does the evidence from transitional job training interventions for ex-offenders support their expansion to other cities? What factors have been found to support positive outcomes in relevant settings?***
- ***To what extent does the evidence of effectiveness for speed enforcement cameras in reducing motor vehicle collisions support including them in suburban traffic safety grant programs?***

2 Identify and gather credible evidence.

In its broadest sense, evidence is "the available body of facts or information indicating whether a belief or proposition is true or valid. As such, evidence can be quantitative or qualitative and may come from a variety of sources, including foundational fact finding (e.g., aggregate indicators, exploratory studies, descriptive statistics, and other research), performance measurement, policy analysis, and program evaluation (see OMB Circular A-11, Section 200.24). Evidence has varying degrees of credibility, and the strongest evidence generally comes from a portfolio of high-quality, credible sources rather than a single study." Because evidence can take many forms and varies in rigor, scope, and specificity, a single study often does not tell us enough information to make a decision or take action. Better answers are found by gathering and synthesizing multiple types and pieces of evidence.

Where can I find evidence?

Key to finding the best available evidence is knowing where to look. Subject matter experts in your agency, including your agency Evaluation Officer, are a valuable resource to help you locate and assess available evidence. They will likely suggest you start with evidence clearinghouses, literature or systematic reviews, academic research databases, and agency websites (for reports and data sets/metrics). evidence.

An evidence clearinghouse is a repository of evidence associated with programs, interventions, and practices that have been reviewed and rated according to criteria. For more information go to https://www.evaluation.gov/resources/#resource=.evidence-clearinghouse&role=*&content=*&year=*

Systematic reviews often grade or classify the rigor of the identified studies to indicate which findings are supported by the strongest evidence.

3 Assess how well the evidence meets your information need.

Having identified and gathered your evidence, you can now consider whether the evidence meets your need. At this stage, you want to assess the **generalizability** of the studies, or the extent to which you can apply research findings based on the sample to the whole population or to other contexts. To understand whether the studies are relevant to your needs, consider the extent to which the studies align with your program or policy, population or outcomes of interest, and context.

Strategies to address generalizability:

- ***Take note of major similarities and differences between the context studied and the context of interest.*** *For example, was the study carried out in geographic regions and with populations that are aligned with your context of interest? When was the study conducted (e.g., is it outdated)? Does the study period reflect an unusual time (e.g., COVID public health emergency) for which findings may not necessarily apply to other contexts?*
- ***Look for sub-group analyses (if available) that might provide more relevant insights.*** *Although many studies focus on presenting key findings of the broad population under investigation, they may include sub-group analyses which might be more applicable to your population of interest. Because these data may not be prominently highlighted, carefully review the methodology and results as well as appendices. In addition, some reports will include language stating that additional analyses can be provided upon request.*
- ***Work through the theory of change to identify where you might learn from the processes even if the impacts can't be assumed to hold up in a new context or with a different population.*** *Each program has an underlying theory of change that describes how and why the given intervention is expected to produce a desired change. Better understanding the necessary steps and causal relationships between inputs, activities, and outputs to achieve desired outcomes may provide insights for comprehensively understanding how change happens that can be applied to your own work*.
- ***Examine results from process evaluations to understand key features of the intervention and its implementation.*** *Although we often want to know whether a program or intervention is effective in producing change, an important complement to outcome evaluations are process evaluations which describe how change happens (or may help explain why it didn't). With its focus on implementation, the types and qualities of services delivered, the resources to deliver those services, and program beneficiaries, process evaluations are valuable for providing insights into how program outcomes were achieved as well as barriers and facilities to achieving those outcomes.*

4 Assess the rigor of the evidence.

Now that you have determined whether the available evidence meets your needs, assess the quality, rigor, and relevance of the studies to determine whether the results – or subset of results – are appropriate for your purpose. The reliability of results varies considerably, due to differences in design, methodology, data collection, and analytic approaches. This diversity in approaches also means that you may encounter conflicting evidence, including null findings (i.e., no statistically significant relationship between groups or variables was found) or inconclusive results when reviewing multiple studies.

To assess the rigor of the evidence and help reconcile conflicting or null findings, consider the strength of the study across key characteristics, as well as its findings. Below are some questions/considerations to ask of the studies under review and their implications for how to apply findings to your work.

Study	Considerations	Implications
Study Design	Was the most rigorous design used for the question being asked? Is the study of high quality?	Relative to other designs, well-conducted randomized controlled trials (RCTs) are considered the "gold standard" approach to demonstrate causality (i.e., to show the impacts of your program or policy). However, some RCTs may be poorly done. Other designs may also be more useful to your needs, depending on your question or problem statement. For example, you may be interested in understanding how a program was implemented, which a process evaluation is better suited to address. **Action**: Although the evidence strength varies by study design, with some designs being stronger than others, it is important to consider how well the study was carried out, regardless of design. Give more weight to findings from high quality studies (e.g., those using representative samples) whose designs directly address your research question.
Sampling Procedures	Does the approach yield a representative sample of the population? For impact evaluations, was an appropriate comparison or control group used?	In quantitative studies, random sampling is preferred over non-random sampling approaches (e.g., convenience sampling, snowball sampling) to reduce the risk of bias in study results. In qualitative studies, however, non-random sampling techniques are often used because the aim is not to test a hypothesis, but to gain an understanding about a population or topic of interest. **Action:** For quantitative studies, give more weight to findings from studies that use a representative, random sample. For qualitative studies, give more weight to findings from studies that draw on a mix of perspectives.

Representativeness	Does the study population align with the population of interest for your program, policy, etc.?	Available evidence may not offer an exact match for the population you are interested in studying. However, some studies may break out findings by subgroups which may more closely reflect your population of interest. **Action**: Give more weight to findings from studies that examine populations that are similar to your population of interest.
Measurement	Did the study use an objective (e.g., test scores) or subjective measure (e.g., self-report of change in knowledge) of the outcome of interest?	A study using pre/post tests to understand the effectiveness of an awareness campaign in changing behavior is stronger when using objective measures than self-reports of intent to change. However, subjective data is not necessarily inferior; the preference for data type depends on the questions. For example, if the campaign is aiming to change attitudes, self-report may be the only way to measure it. **Action**: In general, give more weight to findings from studies that use objective measures. Also keep in mind the limitations of available evidence and its use. Although self-reported data may be more biased than objective data, it may be the only data available and may be more suited to your research question.
Statistical significance	Are the findings statistically significant?	Statistical significance is a way of quantifying whether the observed effect or relationship is likely due to random chance, or attributable to some other factor of interest. The level of significance, usually expressed as a "p-value", refers to the likelihood that the change you observe (relative to no change) is due to chance. A p-value of .05% is the commonly used criterion in studies as the threshold for claims that the observed difference or change was not due to chance. **Action**: Give more weight to findings from studies that report p-values of at least .05. Studies that report p-values lower than .05% provide an even stronger case for relationships that are found.
Practical significance	Are differences that are statistically significant also practically meaningful?	A study may find a 2% difference in outcomes between treatment and control groups to be statistically significant, but the depending on the study, the difference may be too small to be practically meaningful. **Action**: • Consider the magnitude of the difference, not only whether it is statistically significant. Sometimes this requires converting the results into more intuitive or practical units of measurement. • Consider any qualitative findings (e.g., data collected from interviews, focus groups, observations, etc.). Do the qualitative findings provide strong support for the quantitative findings? • Be cautious about over-interpreting results that are seen as "bordering on" or "approaching"

		significance. In other words, is the level of significance close to, but not quite, the desirable level of significance generally used for that field of study?

If conflicting studies are equally rigorous, look for factors related to the intervention that may help explain divergent findings. This may include:

- Program's underlying theory of change
    - Is the program's theory of change similar across the studies? Or are there key differences (e.g., intervention setting such as a school versus a community center) that may have affected outcomes?
- Fidelity or faithfulness to the intervention
    - Was the intervention implemented as intended? Or were changes made that may have affected outcomes (e.g., reduced intervention period to address resource constraints)?
- Populations or contexts
    - Are the studies examining similar populations or contexts? Can divergent findings be explained by differences in the population or contexts being examined (e.g., urban versus rural settings)?
- Presence of specific (unique) factors
    - Did any of the studies examine the intervention under specific circumstances (e.g., economic downturn, election year) that may have affected findings?

All evidence is imperfect in some way. In general, you want to assess the strengths and weaknesses of the design and understand how the specific limitations of the studies under review affect conclusions.

Apply evidence to your work and plan for new evidence needs.

In the final step, synthesize the best available evidence, weighing the rigor and applicability of the evidence at hand and reconciling inconsistencies and gaps where possible to help inform your decision-making. More often than not, you will never find exactly what you need with the evidence available, but you can make the most of what you do find and apply it to your work. Below are ways to apply the evidence you've gathered or to determine new evidence needs to support future work:

- **Apply evidence where consistency is found across studies.**
    - If the contributions of multiple studies (using varied study designs) all point to the same finding, you can confidently apply those findings to your work.
- **Apply evidence in a way that is clear about the study's limitations.**
    - ***Be selective in use of evidence***. Give priority to the highest quality evidence, whether the evidence is drawn from an RCT, outcome evaluation, process evaluation, or other design.
    - ***Limit evidence to your population of interest***. Whereas some studies may focus on the broad population, your population of interest may focus on subgroups. At the same time, you may be able to extrapolate from another population from yours insofar as they share similar characteristics (e.g., youth in foster care to homeless youth).
    - ***Do not stretch findings farther than is appropriate***. For example, are findings limited to specific populations or contexts? To what extent can findings from experimental studies translate to the real world? Are findings mostly based on self-reported data? It should be noted that although there may be biases in self-reported data, these data still add valuable insights, especially when no other data are available or difficult to collect (e.g., elder abuse).
    - ***Make appropriate caveats***. Be explicit about what the data can and cannot say about your population or topic of interest or potential biases in the data. What study design types have been used to examine this policy or program? Are they limited to studies of effectiveness (e.g., outcome evaluations) or studies about implementation (e.g., process evaluations)?
- **Identify gaps where new information/evidence is needed and consider how to build it.** Ask how we can improve on the current evidence. In assessing the best available evidence, what information is missing?