Criteria for Selection of High-Performing Indicators
A Checklist to Inform Monitoring and Evaluation

Goldie MacDonald
Centers for Disease Control and Prevention
Atlanta, Georgia
gim2@cdc.gov

The checklist includes practice-based criteria to be considered in the selection of indicators for use in monitoring and evaluation. The selection of indicators can be a complex, time-consuming task. In some cases, this process is not made explicit for stakeholders. Moreover, those expected to participate in this work come to the discussion with varying levels of knowledge relevant to monitoring and evaluation. Therefore, how do we assess the quality of indicators proposed for use? And, how do we encourage full participation of stakeholders in this dialogue? The purpose of the checklist is three-fold: (1) aid in establishing a process and shared vocabulary for dialogue with stakeholders regarding the selection of indicators; (2) reinforce the necessary connection of indicators to the evaluation questions to be addressed by the study; and (3) contribute to design of data collection activities more clearly linked to intended uses of findings.

As a starting point, what is an indicator? The term is used widely with variation in meaning and application. For the purposes of this checklist, an indicator is a documentable or measureable piece of information regarding some aspect of the program in question (e.g., characteristics of the program, facets of implementation or service delivery, outcomes). In many cases, indicators provide a meaningful marker or approximation of the status of program implementation or outcomes. For the purposes of monitoring and evaluation, an indicator requires an operational definition and methodologically sound, rigorous data collection. An indicator may use qualitative or quantitative information.

The checklist is designed to help those responsible for monitoring and evaluation identify high-performing, resource-efficient indicators in collaboration with stakeholders, especially those in a position to make decisions regarding the program based on findings of the study. The checklist should be used at the earliest stages of planning the study to inform and stimulate dialogue regarding options for indicators, including the practical considerations relevant to data collection.

The criteria that follow are presented in alphabetical order, not in order of importance; the relevance and utility of each criterion will vary by setting and user. This tool does not include all criteria that could be considered when selecting indicators.

☐ Accepted Practice and History of Use

The degree to which use of an indicator is consistent with current and previous practices

In some cases, we select an indicator because it has been used previously. This criterion requires identification of the advantages and limitations of an indicator based on previous use, including use in other contexts or settings; this information should be used to avoid selecting an indicator solely because it has been used in the past. Like many of the criteria to follow, discussion of this criterion may result in not using an indicator. Discussion of accepted practice and history of use of an indicator may result in data collection in a different or novel direction (e.g., to address persistent gaps in knowledge regarding the program, improve the precision or sensitivity of the indicator).

☐ Applicability in Different Settings

The degree to which an indicator is relevant in diverse settings

This criterion is especially important in planning for data collection in multiple locations or sites (e.g., cluster or multisite evaluation). In many cases, the ability to collect comparable and consistent data across sites is a key consideration relevant to the utility of an indicator.

☐ Availability of Data

The degree to which data are accessible for use as part of the study

This criterion requires identification of potential sources of data (primary or secondary) for each indicator. The discussion should include the availability of data at a single point in time (e.g., baseline data) and over time, as appropriate. However, an indicator should not be used solely because data are accessible.

☐ Burden of Data Collection on Participants

The degree to which data collection imposes burden on participants

An assessment of the burden of data collection (human and fiscal) should be completed before agreeing to use an indicator or set of indicators. When planning for monitoring or evaluation in the context of overlapping or related programs in a location or setting, the discussion should include opportunities for collaboration or shared data collection among relevant donors or organizations (to both limit the burden of data collection on participants, and maximize limited resources for the study).

☐ Clarity of Focus and Meaning

The degree to which a single indicator is unambiguous and reflects or represents the evaluand accurately

An indicator should not be used if it will distort or misrepresent the evaluand (i.e., the focus or object of the evaluation), including the conditions or context in which the program is implemented. Stakeholders should view the indicator as a meaningful descriptor or marker of the status of the evaluand. In addition, stakeholders should consider whether or not additional information is required to appropriately or fully interpret data for a specific indicator.

☐ Cultural Appropriateness and Relevance

The degree to which an indicator is culturally appropriate in terms of content or focus and related data collection activities

Stakeholders should determine whether or not the indicator (i.e., the construct and method of data collection) is appropriate and relevant in the context or setting. Typically, these determinations require in-depth, meaningful dialogue with stakeholders.

☐ Data Quality

The degree to which information collected will be complete, reliable, and valid

Thresholds for data quality will vary by setting, based on stakeholder expectations and values. The relative importance of different dimensions or types of quality will also vary. Those responsible for the study should engage stakeholders in establishing explicit standards for data quality.

☐ Investment of Resources

The amount of resources (e.g., funds, personnel, time) needed for data collection, analysis, and use of data or findings

Selection of an indicator requires a precise understanding of the resources needed to collect and analyze the data. This criterion includes consideration of the knowledge or skills necessary to use an indicator. In many cases, discussion of investment of resources results in identification of hard choices to be made regarding feasibility, quality, and timeliness of the data.

☐ Nondirectional Language[1]

The indicator is written as neutral, not defined as positive or negative in advance of data collection

Indicators should not be written with a bias or preference in direction or value. In many cases, program outcomes are presented as statements of achievement or progress (e.g., increase or decrease). Indicators should be written as neutral and nondirectional; the interpretation or valuation of the data for the indicator should occur when data collection is complete. For example, the indicator should be written as “level of knowledge” versus “increase in knowledge” or “prevalence of tobacco use among young people” versus “decreased prevalence of tobacco use among young people.”

☐ Opportunity to Detect Unexpected or Unintended Findings

The degree to which an indicator (or set of indicators) allows for documentation of unexpected or unintended aspects of the program

In many cases, we construct indicators to determine whether or not what we thought would happen actually occurs, both in terms of program implementation and outcomes. However, some of the indicators should be flexible enough to create an opportunity for unexpected or unintended findings to emerge.

☐ Pathway for Use of Data

The degree to which use and users of data are known and agreed upon

This criterion speaks to two issues: who will use data collected and how the information will be used. The criterion is intended to assure that dialogue regarding potential indicators includes explicit and purposeful links to intended use. Accordingly, it is critical that data to be collected are well-suited to the intended use (i.e., provide the types and quality of information needed to achieve the desired use).

☐ Relevance to Evaluation Questions

The degree to which an indicator helps to address predefined evaluation questions

This criterion should be used to assess the connectedness of indicators to the questions to be addressed by the study. If an indicator does not clearly contribute to answering these questions, stakeholders should carefully consider whether or not to use the indicator. In addition, stakeholders should consider whether or not there is sufficient diversity of indicators to adequately address the evaluation questions. The set of indicators may be too focused on a single aspect or dimension of the evaluand (e.g., use of fiscal resources, completion of activities, delivery of programs or services). For example, if stakeholders want to document progress toward intended outcomes over time, the indicators must not be limited to implementation or outputs of the program.

☐ Strength of Evidence or Substantive Merit[2]

The scope and quality of information supporting the indicator as an appropriate descriptor or measure of the evaluand

The degree to which an indicator is considered to be technically sound in a particular domain or field. The evidence for use of an indicator may include published or unpublished literature, stakeholders' experience or values, consultation with those most familiar with the program, or lessons from monitoring or evaluation of similar programs. Stakeholders should determine whether or not the evidence is sufficient to support use of the indicator.

☐ Value within a Set of Indicators

The degree to which a single indicator adds meaning to a set of indicators

In many cases, stakeholders consider an indicator as a single item or piece of information, as opposed to a component of a set of indicators intended to address predefined monitoring or evaluation questions in concert. In addition to assessing the individual merits of an indicator, stakeholders should assess the value of an indicator within a set of indicators to fully determine its utility. This criterion aids in identifying duplications or redundancies in data collection and may result in use of a more manageable, smaller number of high-performing indicators.

References

[1] Bloom, S. S. (2008). *Violence against women and girls: A compendium of monitoring and evaluation indicators.* United States Agency for International Development. Retrieved April 18, 2012 from www.awid.org/Library/Violence-against-Women-and-Girls-a-Compendium-of-Monitoring-and-Evaluation-Indicators

[2] Indicators Technical Working Group of the UNAIDS Monitoring and Evaluation Reference Group. (2010). *Indicator standards: Operational guidelines for selecting indicators for the HIV response.* Joint United Nations Program on HIV/AIDS. Retrieved March 28, 2013 from http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/4_3_MERG_Indicator_Standards.pdf

Acknowledgments

The author acknowledges the thoughtful discussion and input of participants at the Australasian Evaluation Society 2011 International Conference, as well as contributions to earlier versions of the content by Danyael Garcia, John Houseal, Laura Lessard, Ann Moen, Janis Weber, Boris Volkov, Lori Wingate, and two anonymous reviewers.