

## Think Tank Summary: Challenges in Measuring Implementation Success

Carol VanDeusen Lukas, EdD  
Carmen Hall, RN, PhD

The measurement of implementation success – that is, the extent to which an intervention is put into practice as intended -- is a critical aspect of implementation science in two ways. First, it is essential as the dependent variable in identifying effective implementation strategies. In order to understand what strategies are or are not successful and to identify the factors that affect implementation, we must be able to determine what constitutes more and less successful implementation. Second, as an independent variable, measurement of implementation success contributes to understanding the success or failure of the intervention outcomes.

A few years ago, relatively few researchers in health care recognized the importance of measuring the success of implementation separately from the effectiveness of the innovation. Today there is greater recognition of the importance documenting the innovation in practice, as illustrated by the number of sessions dealing with implementation fidelity in this conference. However, there is no consensus on conceptual frameworks, shared definitions or standards for success, despite several recent attempts to present frameworks. With everyone using somewhat different terms, it is difficult to compare frameworks and difficult to compare study findings.

This difficulty is underscored by the findings of a recent inventory of implementation studies in the Health Services Research and Development Service in the Department of Veterans Affairs. While a substantial majority of studies inventoried included some indicators that could be considered measures of implementation success, the measures were not consistently well developed. Even collecting information about measurement was difficult because there was no shared language for talking about these concepts and measures. Across studies, investigators both used the same terms in different ways and applied different labels to the same constructs. In some cases, after probing, the team conducting the inventory identified measures in a project that they categorized as measures of implementation success even though the study PI did not label them – or in some cases think of them -- in those terms.

Thus, one starting place for advancing this aspect of implementation science is to create a shared framework of constructs and definitions for measuring implementation success. The first objective of this think tank session was to stimulate development of that framework. The second objective was to discuss the measurement of these constructs. The session was organized around questions that we introduced and illustrated with specific examples from our own research.

### 1. Creating a shared frame of reference

To begin the discussion of shared terminology, we proposed as a frame of reference a basic and highly simplified logic model for implementing an innovative practice in a healthcare setting. The model posits that *implementation strategies* will lead to *use of the clinical innovation in practice* which in turn will lead to the desired *process outcomes* which will result in *improved health outcomes*. Further, all aspects of the process and outcomes will be influenced by a wide range of *other factors* in the context in which the innovation is being implemented. While these contextual factors are critical to the change process and implementation success, they are not directly involved in measuring implementation success in the sense of measuring the *intervention* in practice – that is, the documentation of the implementation strategies that are actually used and the clinical innovation as it is put into practice. The intent of the logic model was not to capture the dynamics of the implementation process with a full elaboration of the range of factors that affect implementation and a portrayal of the multiple iterations and feedback loops involved in complex change processes. Instead its purpose was to focus discussion on the components of the model and their definitions. Specifically, we wanted to highlight the distinctions between implementation strategies, the clinical innovation in use and the desired clinical outcomes, and to offer definitions for discussion:

- **Clinical innovation** refers to the new practice or technology that is to be implemented. An innovation is defined in terms of being used in an organization for the first time. An innovation may be specifically focused, such as hand-hygiene compliance or clinical guidelines, or broadly, such as a

collaborative care model for bipolar disorder. It may be targeted to the individual practitioner, to the microsystem in which practitioners work and/or to higher levels of the organization.

- **Implementation strategies** are the processes used and activities undertaken to introduce and support the innovation as it is brought into regular use. They may follow an established implementation approach (such as external facilitation or Replicating Effective Programs (REP) or one that is locally developed. Some projects do not use an explicit implementation approach but still need in some way to take steps to bring the innovation to the organization.
- **Clinical innovation in use** is the innovation as it is put into routine practice by staff in the organization. This is the objective toward which the implementation strategies are directed. In the logic model, the clinical innovation in use serves both the outcome of the implementation strategies and a determinant of the desired clinical outcomes. Depending on the levels of the organization targeted by the innovation, its use will affect individual practitioners, microsystems and/or larger components of the organization.
- **Intervention** refers to the combined implementation strategy and clinical innovation being introduced in a study. In traditional health services research where the intervention is under the control of the research team, intervention usually refers only to the clinical innovation that is being put in place. In implementation research, where implementation moves into practice settings beyond direct control of researchers and therefore where others responsible for implementation, the implementation processes and activities also are part of the intervention. Depending on the study focus, implementation strategies may receive more or less attention. In studies that are formally testing an implementation strategy, this aspect of the intervention will be prominent. In studies that do not use an explicit implementation approach, implementation strategies may not be recognized as a major part of the study, but it should be recognized that something is being done to bring the innovation into the organization.
- Desired **clinical process outcomes** are the changes in practice that are expected to lead to the desired health outcomes. They are the intermediate results of the innovation in use, for example, adherence to clinical practice guidelines.
- Desired **health outcomes** are the improved health status of a cohort or population of patients (such as low infection rates, decreased morbidity or increased quality of life years) that are sought by implementing the innovation. In implementation research, it is expected that the outcomes are indicators that have been demonstrated effective in prior research. Hence, the implementation study may or may not measure these outcomes independently – investigators may judge the evidence linking presence of the innovative practice to the outcome to be strong enough that it does not have to be tested again, and they instead invest the study resources in implementation. Alternatively, for an early implementation study or one testing an innovation in a new setting or with a different population, investigators may decide they need to directly test effectiveness to make sure it holds up under the new conditions. Hybrid studies that test both effectiveness and implementation are gaining attention.

Despite the stated intention of using the logic model simply as a frame of reference for developing shared definitions of its components, it was difficult for participants to move beyond the model's static and highly simplified presentation. The discussion focused on the complexity and dynamics the implementation processes. Objections were not to the component definitions, but to 1) the absence of feedback loops and multiple iterations in the logic model that would more accurately portray the implementation process, and 2) the lack of elaboration on the myriad of factors affecting the process and attention to unintended consequences. The discussion thus highlighted an important challenge in creating a shared framework in this focused aspect of implementation science: the difficulty of differentiating the dynamics of the implementation process from the more focused construct of the intervention in practice. One proposed solution was to define the component terms without portraying them in a logic model.

## 2. Span and standards for measuring implementation success

A second set of key challenges in measuring implementation success lies in operationalizing the definition of success. Implementation studies vary widely in terms of type of innovation, implementation

strategy, and implementation setting. Therefore, an expectation of a single standard or measure is not appropriate. At the same time, this diversity makes a shared framework for the span of and standards for measurement even more important.

The measurement of implementation success is generally defined as the presence of the innovation as intended in practice. However, following the earlier argument that the intervention includes both implementation strategies and the clinical innovation, we proposed that, depending on the structure of the study and types of questions it is designed to address, measures of implementation success potentially should document the implementation strategies used as well as the clinical innovation in practice. If a study is testing a specific implementation method, such as Replicating Effective Programs, it will be important to determine what was actually done and whether the method was used as intended. Even without testing a formal implementation strategy, investigators should document their activities in introducing and supporting the innovation. In both cases, the information will be relevant to interpreting study findings, and if successful, spreading the innovation to other settings.

Another critical aspect of measuring implementation success is to identify a range of explicit standards for success. The choice of standards and the levels at which the success of implementation will be judged for a particular study will depend on the intervention and the aims of the study. However, by agreeing on definitions for a range of standards – and expecting implementation study protocols to be explicit about which are being used – we will come closer to being able to compare findings across studies.

Fidelity is the most commonly-used standard of success. Traditionally, fidelity has been defined in terms of adherence to the intent of the innovation developers. More recently it has been defined more comprehensively to include not only adherence, but also scope/reach/reliability, intensity/dose, quality of program, participant responsiveness, sustainability/maintenance. For example, scope/reach/reliability reflects the spread of the innovation with scope or reach indicating the range of application of the innovation in terms of numbers of staff and locations and reliability indicating the consistency of its use. Intensity or dose reflects the quantity and depth of actions taken to implement the innovation, indicating the amount of leverage over outcomes that a given action could be anticipated to have.

In the discussion, there was general agreement with the expanded standards, although as pointed out under point 3 below, there was some question about the appropriateness of the use of the term fidelity. Issues for continuing discussion and for practical development measures in studies include:

- Are all of these standards appropriate in all situations?
- Are there other standards against which implementation should be judged?
- How do investigators determine the appropriate standards for a specific study?

### **3. *Defining a complex innovation in operational terms***

Another important challenge in creating measures of implementation success is defining operationally the innovation that is being implemented. With complex innovations involving interdependent individuals and parts of the organization, this is both critical and often difficult. In traditional research designs, the expectation is that the innovation will be fully specified in advance of the study and held constant in a controlled setting for the study duration. But in implementation research, where the whole purpose is to introduce the innovation to ongoing operations, strict detailed replication is neither realistic nor desirable. Consistent with theories of organizational change, innovations are more likely to be successfully implemented and sustained if they can be tailored or adapted by the people using them to fit their organization. Therefore, some variation across users and settings is to be expected. At the same time, there must be boundaries to the variation. Following from this premise, a key challenge both to the success of implementation and to the development of a measure of implementation success is to define the features that are core to the innovation. Such definition is necessary to distinguishing tailoring – where the core elements are present but have been put into practice in ways that best fit that organization -- from partial implementation.

In the discussion, there was agreement that local adaptation and tailoring of an innovation to specific settings are critical. Hence, in measuring implementation success in any given study, investigators and their local colleagues responsible for implementation will need to address several questions:

- Which features have to be present for the innovation to maintain its effectiveness?
- Which can vary without compromising effectiveness?
- How are those core features defined in operational terms so that both participants and research team recognize their presence or absence?

In the context of tailoring and adaptation, the discussion also raised questions about whether fidelity to the intended innovation will always be an appropriate standard or whether a more evolutionary concept of success might be needed. This is an issue that deserves further discussion.

#### **4. *Identifying appropriate measurement approaches***

As in most aspects of research, there are a variety of data collection approaches and tools that can be used to measure implementation success. The appropriateness of the choice in any given study will depend on both: 1) which measures will give the best and most rigorous answers to the type of intervention and questions the study is aiming to address and 2) what is feasible given the scope of the project and the resources available for data collection. Most projects face constraints. Among the issues to be considered in developing data collection strategies within these constraints:

- At what levels of the organization is implementation targeted and is data collection needed?
- Are secondary data sources available that tap the indicators sought?
- If primary data will be collected, will it be through surveys that cut widely across the organization or through interviews and focus groups that probe more deeply?
- Are validated data collection tools available and appropriate?
- What data collection will be expected of local participants and what will be done by the study team?

The discussion highlighted the importance of differentiating levels of the intervention and thus developing appropriate levels of measurement.