Advancing Evidence Synthesis from Effectiveness to Implementation: Integration of Implementation Measures into Evidence Reviews

Aaron Tierney  
*Stanford University School of Medicine*

Marie Haverfield  
*San Jose State University, marie.haverfield@sjsu.edu*

Mark McGovern  
*VA Health Care System*

Donna Zulman  
*Stanford University School of Medicine*

Follow this and additional works at: https://scholarworks.sjsu.edu/comm_pub

Part of the Communication Commons

**Recommended Citation**


This Article is brought to you for free and open access by the Communication Studies at SJSU ScholarWorks. It has been accepted for inclusion in Faculty Publications by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.
health services outcomes (quality of life, cost of care, etc.), an approach that does not provide the information necessary to understand potential implementation challenges and opportunities. The findings from these synthesis efforts often inform the design of future interventions, and this gap creates a cycle where important implementation considerations are ignored, preventing interventions from moving beyond an initial pilot stage. Incorporating implementation and dissemination measures into evidence synthesis efforts could help facilitate the spread of high-value, effective, and sustainable interventions.

The objective of this paper is to evaluate the feasibility of extracting factors related to implementation during evidence synthesis in order to enhance the replicability of successes of studies of interventions in health care settings. We describe opportunities to use established measures from the field of implementation science to abstract information from health care delivery intervention studies (clinical studies in which participants are assigned to groups that receive one or more intervention/treatment [or no intervention] so that researchers can evaluate the effects of the interventions on biomedical or health-related outcomes). We illustrate the application of these recommendations by presenting a systematic review of clinician-patient communication skills interventions.

**DEFINING TERMINOLOGY**

We reviewed landmark papers describing implementation measures to identify widely used appropriate and feasible constructs to abstract during literature reviews. Our goal in selecting measures was to codify previous interventions’ technical components and underlying social, contextual, and dynamic factors that may affect implementation, which are crucial to successful replication. Table 1 presents 10 identified factors: acceptability*, adoption*, appropriateness*, feasibility*, fidelity*, implementation cost*, intervention complexity, penetration*, reach, and sustainability*. Most of these factors (marked with an asterisk above) are outlined and defined in Proctor et al. The constructs of reach and adoption derive from the Reach Effectiveness-Adoption Implementation Maintenance (RE-AIM) model, a widely used framework for evaluating implementation outcomes. In addition, we included intervention complexity, a core construct in the Consolidated Framework for Implementation Research (CFIR) and other widely used implementation science frameworks.

**IMPLEMENTATION FACTORS AND ABSTRACTION RECOMMENDATIONS**

Below we describe strategies for abstracting ten implementation factors from intervention studies during evidence synthesis. In cases where included studies cite previous papers that describe the intervention, these papers should also be reviewed. Table 1 presents a definition of each factor, examples of relevant information that can be abstracted from intervention studies, and the sections in which the relevant information is most often found.

**Acceptability.** During implementation, acceptability is the “perception among implementation stakeholders that a given treatment, service, practice, or innovation is agreeable, palatable, or satisfactory.” Acceptability of an intervention by the target population is strongly associated with likelihood of adoption, as described in seminal work by Everett Rogers. Acceptability is typically measured through a survey of stakeholders and often focuses on provider and/or patient satisfaction. If acceptability is not assessed in an intervention trial, it might be inferred in part from retention of providers and patients throughout the study, as dropout rates may indicate that participants do not find the intervention to be worthwhile, satisfactory, or agreeable. Caution should be taken in extrapolating retention rates to acceptability, because of the multiple factors that influence study retention, and comparing retention rates across studies that place varying emphasis on retention (e.g., highly controlled studies vs. pragmatic studies) could be misleading. However, if dropouts are surveyed or interviewed, these reported findings might provide a source of information about acceptability.

**Adoption.** Adoption is the “intention, initial decision, or action to try or employ an innovation or evidence-based practice.” Adoption is often influenced by other implementation factors (e.g., acceptability, feasibility, implementation cost) and is an outcome measure of implementation. Adoption can be extrapolated from the proportion of eligible providers (or clinics) that participate in part or all of an intervention.

** Appropriateness.** Appropriateness (sometimes referred to as compatibility) is the “perceived fit, relevance, or compatibility of an innovation or evidence-based practice for a given practice setting, provider, consumer, or problem.” Interventions perceived as more appropriate by stakeholders are more likely to be adopted. Appropriateness can be determined through provider satisfaction surveys that measure provider opinions about intervention usefulness or impact on workflow, as well as pre-implementation considerations about fit (i.e., Does the hospital have the proper resources to be able to implement the intervention? Does the intervention make sense with the population of a given setting?).

**Feasibility.** Feasibility is the “extent to which a new treatment, or an innovation, can be successfully used or carried out within a given agency or setting.” Settings with adequate time and resources (both actual and perceived) to adopt and continue an intervention are more likely to initiate the use of that intervention. Feasibility considerations mostly consist of demands put on a provider or a system, including the cost of the intervention and the time and staffing demands to properly implement the intervention. Papers may address feasibility by
describing the training time and staffing required, start-up costs, materials needed for the intervention, and the need for systems-level changes for implementation.

**Fidelity.** Fidelity is the “degree to which an intervention was implemented as it was prescribed in the original protocol or as it was intended by the program developers.” Higher levels of fidelity are associated with greater improvement in clinical outcomes. Fidelity is most relevant when an intervention is tested in multiple trials. It is typically measured through a fidelity checklist or post-implementation assessment of the degree to which the intervention was implemented as planned. In an evidence synthesis project, fidelity can be evaluated by examining multiple studies of the same intervention to determine whether later versions of the intervention resemble the original intervention or whether adaptations were made (and if so, the reasons why).

**Implementation Cost.** Implementation cost is the “cost impact of an implementation effort” and is often influenced by intervention complexity, implementation strategy, and setting. High implementation costs are frequently a barrier to adoption. Implementation costs are specific to costs related to the implementation process and are distinct from cost (or cost-effectiveness) outcomes that result from an intervention. Implementation costs are rarely reported in efficacy studies but may be estimated by reviewing reports of personnel requirements (including the need for new staff, or time demands for training/participation) and technology or equipment.

**Intervention Complexity.** Intervention complexity is the “perceived difficulty of implementation, reflected by duration, scope, radicalness, disruptiveness, centrality, and intricacy and number of steps required to implement.” Complexity can also influence the reproducibility of an intervention and the external validity of trials the intervention undergoes. Intervention complexity can be assessed by examining the time period over which an intervention is carried out, the number of steps and personnel required to carry out the intervention, or the number of trainings required to teach an intervention.

**Penetration.** Penetration is the “integration of a practice within a service setting and its subsystems.” Interventions that have buy-in from an organization and/or greater organizational support have greater uptake. Penetration shares elements with reach, but is a system-level measure of the quality and scope of adoption and is defined by the proportion of providers who use an intervention out of the anticipated providers. Penetration can be observed through multi-site data (if available) or through a description of intervention integration within a system.
Reach. Reach examines “the absolute number, proportion, and representativeness of individuals who are willing to participate in a given initiative [or intervention].”9 The greater the population that an efficacious intervention reaches, the greater positive impact it will have on population health outcomes. Certain measures of reach could also ensure that underserved populations receive appropriate care and services. Demographics tables typically present the characteristics of the population that received an intervention. In some cases, flow diagrams may present information about the characteristics of patients who dropped out of a study. Combining these data with information about the target population can provide insight about the representativeness of the study population and the patients who received the full intervention.

Sustainability. Sustainability (or maintenance)9 is the “extent to which a newly implemented treatment is maintained or institutionalized within a service setting’s ongoing, stable operations.”10 Organizations with resources to sustain an intervention are more likely to adopt and continue an intervention.11 An intervention that is more easily sustainable will also have greater longevity. Sustainability is rarely described in early intervention studies, but the discussion section may consider characteristics of the intervention or implementation process that influence sustainability. In an evidence synthesis, follow-up studies or reports may provide information about whether an intervention was expanded or sustained after initial financial and workforce support ends.

SYSTEMATIC REVIEW

Methods

To illustrate the feasibility of abstracting the above implementation measures, we conducted an evidence synthesis of 29 studies of physician communication skills interventions published between 1997 and 2017 (see Appendix). To be included, studies had to be randomized controlled trials or controlled observational studies focused on an interpersonal intervention aimed at improving clinician communication skills, and had to report an outcome aligned with the Quadruple Aim of improving patient experience, improving provider experience, containing costs, and improving population health outcomes.21 The studies were a subset of articles from a systematic review that was conducted for a previous evidence synthesis project.22 For this review, we focused on studies of physician communication skills interventions to facilitate comparison of implementation factors across similar types of interventions.

For this proof-of-concept systematic review, our objective was to determine whether implementation factors could be reliably abstracted using the above guidelines. Two reviewers (AT, GP) independently abstracted data about implementation considerations from each article. For each of the 10 aforementioned factors, the reviewers determined whether the measure was described in the article text, tables, appendices, and (when cited) previous publications about the intervention. In order for a “measure to be described,” both reviewers needed to identify supporting information in the paper. Discrepancies were resolved by consensus with a third reviewer (MH). Percent agreement was initially 69%. However, one of the raters was trained in implementation science methods and the other rater was new to the material. Consensus ratings favored the reviewer with more implementation science experience in 70 of the 91 (77%) disagreements.

Results

Among the 29 studies, there was variation in the proportion that reported information relevant to acceptability (48%), adoption (76%), appropriateness (66%), feasibility (62%), fidelity (66%), implementation cost (79%), intervention complexity (100%), penetration (76%), reach (86%), and sustainability (24%). Approximately two-thirds (19/29) of studies included a measure of fidelity. Only 1 (3%) study contained information related to all 10 of the implementation measures. Detailed data on the studies and their considerations can be found in Table 2 and Appendix Table. Based on these results, the full systematic literature review incorporated a measure of intervention complexity (given that it was widely addressed). In doing so, the review revealed that interpersonal communication interventions with lower demands on provider time and effort were often as effective as those with higher demands.22

DISCUSSION

To fully understand the downstream potential of an intervention, it is important to consider its effectiveness and the process, demands, and resource requirements involved in implementation. While there is variation in the degree to which intervention studies currently report implementation factors, there is a growing demand for this information. As is evident from our systematic review, it is often feasible to extract implementation information during a literature review even though it might not be explicitly reported in analyzed studies. As intervention studies increasingly report implementation factors, evidence synthesis efforts should evolve to abstract and synthesize this information.

We illustrate that abstracting information about factors such as acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, implementation complexity, penetration, reach, and sustainability may enhance the value of systematic reviews. Including this information will inform clinicians, researchers, and health system leaders about stakeholder preferences, resource demands, and other factors that may influence implementation and dissemination success. In the instance of our proof-of-concept extraction, we found that all papers included some description of intervention complexity,
Table 2 Number of systematic review trials with information on each implementation consideration (full references for systematic review trials are available in Appendix)

<table>
<thead>
<tr>
<th>Implementation factor</th>
<th>Trials reporting relevant information, n (%)</th>
<th>Percent agreement regarding presence of factor (%)</th>
<th>Examples of information abstracted from articles</th>
<th>Relevance to implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>14 (48%)</td>
<td>48%</td>
<td>“Some team members did not agree to participate in the training program (as a result of a lack of time or interest, leaving the team between the assessment and training periods, concerns with role play, or being already involved in another training).” (Merckaert 2015)</td>
<td>The intervention may take too much time or not be considered worthwhile by relevant parties, so interventions with similar characteristics may face the same barriers to implementation</td>
</tr>
<tr>
<td>Adoption</td>
<td>22 (76%)</td>
<td>79%</td>
<td>“Thirty (19.6%) of 153 oncologist candidates returned the consent form, whereas the others chose not to participate.” (Hart 2006)</td>
<td>Low adoption may mean that individuals trying similar interventions should think about different design approaches or ways to decrease demand placed on clinicians in order to increase adoption</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>19 (66%)</td>
<td>66%</td>
<td>“We chose for our training program teaching objectives which on the one hand represent the most frequently studied skills and on the other hand seemed to us specific for the needs of patients in obstetrics and gynaecology.” (Alder 2007)</td>
<td>These researchers focus on including items in their training that would have the most utility for the audience they were trying to reach. This could have impact on other factors as well, including adoption and acceptability</td>
</tr>
<tr>
<td>Feasibility</td>
<td>18 (62%)</td>
<td>69%</td>
<td>“Trainees further had access to an online training support portal and an online course for those who were not able to attend all live sessions.” (Jjam 2017)</td>
<td>Offering an online training course allows physicians to participate on their own time, rather than having to force a training into their already busy schedules there by facilitating participation</td>
</tr>
<tr>
<td>Fidelity</td>
<td>19 (66%)</td>
<td>59%</td>
<td>“For the experimental group, interviews were all audio-taped and transcribed. The transcripts were then assessed by two trained investigators, in order to determine the number of communication skills used within the six-step SPIKES.” (Ajam 2017)</td>
<td>Even if a trial of an intervention has a positive effect on outcomes, it is important to verify that it is being delivered as intended. This may lead to modifications or to explanations of why an intervention did not behave as planned</td>
</tr>
<tr>
<td>Implementation cost</td>
<td>23 (79%)</td>
<td>76%</td>
<td>“Training was carried out by a national trainer with experience in communication skills together with members of the research team who observed and facilitated the training. In total, nine training workshops were conducted at the Faculty of Medicine, Damascus University. The duration of each workshop was 20 h in total, delivered over 3 days.” (Bashour 2013)</td>
<td>If the intervention requires special training or certain staff such as someone trained deeply in communication skills, this could create a cost barrier to either research teams or clinics to implementing an intervention</td>
</tr>
<tr>
<td>Intervention complexity</td>
<td>29 (100%)</td>
<td>79%</td>
<td>“The trained sequence of Establishing Focus (EF) skills included: 1) orienting the patient to the EF process, 2) asking the patient to list concerns, 3) making space for pressing patient stories early on, when necessary, 4) avoiding premature ‘diving’ into diagnostic sequences or patient story telling before a full agenda has been set, 5) asking the patient to prioritize their concerns, 6) when necessary, negotiating priorities with the patient, and 7) seeking confirmation and commitment from the patient.” (Brock 2011)</td>
<td>A more complex intervention may be less likely to be adopted and complexity can influence fidelity if the intervention is too hard to remember or has too many steps involved. Also, extracting this type of information may give researchers examples of how to effectively implement similar interventions and may aid in design</td>
</tr>
<tr>
<td>Penetration</td>
<td>22 (76%)</td>
<td>76%</td>
<td>“Participation in the study was mandatory for all senior physicians and residents employed at the departments of internal medicine.” (Maatouk-Bürmann 2016)</td>
<td>Organizational support and integration can have effects on other implementation factors, both positively and negatively. Seeing how an intervention behaves with or without organization support or integration can be informative for dissemination and implementation efforts of similar interventions</td>
</tr>
<tr>
<td>Reach</td>
<td>25 (86%)</td>
<td>83%</td>
<td>An intervention designed to improve patients’ self-perceived health was used</td>
<td>If an invention is tested with a diverse subset of patients, it may be more</td>
</tr>
</tbody>
</table>

(continued on next page)
and this information could be used to examine the association between intervention complexity and study outcomes. Such information could also provide justification for certain implementation strategies or inspire innovations to address implementation barriers. Standardizing the practice of incorporating implementation information could also improve intervention design methods, resulting in greater uptake and implementation of interventions with proven efficacy on improving outcomes of interest.

We found few implementation measures are routinely reported or described, although several can be estimated from frequently reported information. Some factors were buried in unusual places in the text, requiring a comprehensive review of each article and potential review of previous publications that describe the intervention under investigation. However, all the studies reported information relevant to multiple implementation factors, although the authors of these studies were likely not actively considering them during authorship” with “might not have explicitly labeled them as such. This demonstrates that our method is appropriate and feasible with a diverse pool of studies. The time and resources required to abstract implementation factors are also likely to vary by topic and intervention complexity.

This study builds on a growing body of literature that describes recommendations for incorporating implementation science into intervention design, evaluation, and evidence synthesis. We advocate for the development and adoption of a more standardized structure for reporting implementation factors in intervention trials to facilitate integration of this information into evidence synthesis. For example, studies of new interventions should consider evaluating stakeholder perceptions (e.g., acceptability, appropriateness, feasibility). One mechanism to support this could be a new checklist for the Consolidated Standards of Reporting Trials (CONSORT), similar to the non-pharmacological treatment intervention extension, which already includes constructs related to appropriateness (tailoring for each participant) and fidelity. CONSORT is widely used in medical literature and includes the creation of the CONSORT statement from checklists that help ensure transparent and complete reporting in randomized trials. The Quality Improvement Minimum Quality Criteria Set (QI-MQCSS), which offers guidelines for evaluating quality improvement interventions, is another resource and incorporates domains such as fidelity, reach, and sustainability. Finally, augmenting the Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) may be an effective avenue to encourage the inclusion of implementation factors in evidence reviews. More extensive adoption of certain existing reporting approaches (such as Standards for Reporting Implementation Studies [StaRI]) could also advance this goal.

Similarly, it is important to publish information about the intervention and implementation processes (in a background paper or an appendix, if necessary) that can form the basis for understanding intervention complexity and implementation costs. Reach can usually be extrapolated from conventional demographics papers, but it is valuable to incorporate a discussion about the representativeness of the study population to the potential population that may benefit from an intervention. Other factors, such as fidelity and sustainability, may need to be reported in follow-up papers. Standardizing and explicitly labeling implementation factors will help eliminate error in the abstraction process and reduce the time and effort needed to abstract and synthesize this information.

Importantly, the goal of this paper is to demonstrate the feasibility of abstracting information about implementation factors. The specific factors we chose will benefit from iteration and refinement, as well as expert review. A possible limitation of abstracting implementation data is the additional personnel time required to perform the extraction. For our systematic review, two reviewers completed implementation factor abstraction in just over 1 week. Another limitation is
that previous experience with implementation science may be an important consideration in accurately identifying implementation factors, particularly when they are not explicitly defined in the text. Not all teams are familiar with implementation science and additional training may be required. Finally, some implementation factors are qualitative in nature or lack standardized measures, so it may not be feasible to synthesize this information in meta-analyses. This may change given heightened interest in quantitative methodology in the field of implementation science. The qualitative nature of the process also creates the possibility of bias and misinterpretation of information presented in studies. Extra care should be taken before making inferences from data garnered from this abstraction method.

In summary, while there is variation in the degree to which intervention studies currently report implementation factors, there is a growing demand for this information. This paper aims to enhance the evidence synthesis process by providing researchers with potential implementation factors to aid in translation. While abstracting information surrounding these 10 factors does not guarantee replicability, it can reveal barriers to implementation. It can also highlight factors that were not addressed in previous studies that future implementers may explore. Recent recommendations by Powell et al. and the CFIR website also stress the importance of tailoring and adapting implementation and sustainment strategies. Abstracting information about these 10 factors could provide initial insight into strategies that may be helpful in similar settings.3, 10 Abstracting information about implementation factors may increase the value of evidence synthesis efforts, improving the dissemination and adoption of interventions that are effective, feasible, and sustainable across different contexts. In addition, enriching reviews with this information is likely to create a new avenue to increase the cumulative knowledge of implementation science among the scientific community, and support design efforts that facilitate dissemination.

Acknowledgments: We would like to thank Gabrielle Pecinovski for her contribution to evidence abstraction in the proof-of-concept systematic review. We would also like to thank Rachel Schwartz, PhD; Michelle B. Bass, PhD, MSL, AHIP; Cati Brown Johnson, PhD; Nadia Safaeezad, MPH; Dani Zions, MScPH; Jonathan G Sheu, MD, MS; Sonsoo Thadaney, MBA; Karl A Lorenz, MD, MS; Steven M. Asch, MD, MPH; Abraham Venghese, MD; Laura Jacobson, MPH; Shreyas Bharadwaj; and Isabella Romero for their contribution to the overall systematic review that laid the initial groundwork for the proof-of-concept systematic review in this paper.

Corresponding Author: Donna Zulman, MD, MS, Division of Primary Care and Population Health, Department of Medicine, Stanford University School of Medicine, Medical School Office Building (MSOB), 1265 Welch Road, MC 5411, Stanford, CA 94305, USA (e-mail: dzulman@stanford.edu).

Funding Information This study was funded by a grant from the Gordon and Betty Moore Foundation (#65382). PIs Donna Zulman and Abraham Venghese. Dr. Haverfield was supported by VA Palo Alto Center for Innovation to Implementation (Ci2i) HSR&D postdoctoral fellowship.

Data availability The datasets during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with Ethical Standards: Conflict of Interest: The authors declare that they do not have a conflict of interest.

Disclaimer: The contents of this article do not represent the views of VA or the United States Government.

REFERENCES


**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.