About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world’s best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org
METHODS AND APPLICATIONS IN IMPLEMENTATION SCIENCE

Topic Editors:
Mary E. Northridge, New York University (NYU) Langone Dental Medicine – Brooklyn, United States
Donna Shelley, NYU School of Medicine, United States
Thomas G. Rundall, University of California, Berkeley School of Public Health, United States
Ross C. Brownson, Washington University School of Medicine, United States

The Trans Adriatic Pipeline depicted here starts from the Caspian Sea and reaches the coast of southern Italy. It is intended as a tribute to Lawrence Green, DrPH, MPH, who popularized the pipeline graphic that depicts the 17-year odyssey necessary for the production and transfer of knowledge from research to practice and policy.

The purpose of this Research Topic is to share the latest developments in the methods and application of implementation science. Briefly, implementation science is the study of methods to promote the adoption and integration of evidence-based practices, interventions, and policies into routine health care and public health settings. Implementation research plays an important role in identifying barriers to, and enablers of, effective health systems programming and policymaking, and then leveraging that knowledge to implement evidence-based innovations into effective delivery approaches.

# Table of Contents

## INTRODUCTION TO THE RESEARCH TOPIC

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>06</td>
<td>Editorial: Methods and Applications in Implementation Science</td>
<td>Mary E. Northridge, Donna Shelley, Thomas G. Rundall and Ross C. Brownson</td>
</tr>
</tbody>
</table>

## THEORIES, FRAMEWORKS, AND APPROACHES

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08</td>
<td>From Classification to Causality: Advancing Understanding of Mechanisms of Change in Implementation Science</td>
<td>Cara C. Lewis, Predrag Klasnja, Byron J. Powell, Aaron R. Lyon, Leah Tuzzio, Salene Jones, Callie Walsh-Bailey and Bryan Weiner</td>
</tr>
<tr>
<td>14</td>
<td>RE-AIM in Clinical, Community, and Corporate Settings: Perspectives, Strategies, and Recommendations to Enhance Public Health Impact</td>
<td>Samantha M. Harden, Matthew Lee Smith, Marcia G. Ory, Renae L. Smith-Ray, Paul A. Estabrooks and Russell E. Glasgow</td>
</tr>
<tr>
<td>24</td>
<td>Enhancing the Impact of Implementation Strategies in Healthcare: A Research Agenda</td>
<td>Byron J. Powell, Maria E. Fernandez, Nathaniel J. Williams, Gregory A. Aarons, Rinad S. Beidas, Cara C. Lewis, Sheena M. McHugh and Bryan J. Weiner</td>
</tr>
<tr>
<td>33</td>
<td>The Price per Prospective Consumer of Providing Therapist Training and Consultation in Seven Evidence-Based Treatments Within a Large Public Behavioral Health System: An Example Cost-Analysis Metric</td>
<td>Kelsie H. Okamura, Courtney L. Benjamin Wolk, Christina D. Kang-Yi, Rebecca Stewart, Ronnie M. Rubin, Shawna Weaver, Arthur C. Evans, Zuleyha Cidav, Rinad S. Beidas and David S. Mandell</td>
</tr>
</tbody>
</table>

## PARTNERSHIPS, ENGAGEMENT, AND COLLABORATION

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>Variation in Research Designs Used to Test the Effectiveness of Dissemination and Implementation Strategies: A Review</td>
<td>Stephanie Mazzucca, Rachel G. Tabak, Meagan Pilar, Alex T. Ramsey, Ana A. Baumann, Emily Kryzer, Ericka M. Lewis, Margaret Padek, Byron J. Powell and Ross C. Brownson</td>
</tr>
<tr>
<td>64</td>
<td>A Protocol for a Feasibility and Acceptability Study of a Participatory, Multi-Level, Dynamic Intervention in Urban Outreach Centers to Improve the Oral Health of Low-Income Chinese Americans</td>
<td>Mary E. Northridge, Sara S. Metcalf, Stella Yi, Qiu-ji Zhang, Xiaoxi Gu, Chau Trinh-Shevrin for the Implementing a Participatory, Multi-Level Intervention to Improve Asian American Health Research Team</td>
</tr>
</tbody>
</table>
84  A Mixed Methods Approach to Evaluate Partnerships and Implementation of the Massachusetts Prevention and Wellness Trust Fund
Rebekka M. Lee, Shoba Ramanadhan, Gina R. Kruse and Charles Deutsch

MEASUREMENT ADVANCES
94  Developing a Survey Tool to Assess Implementation of Evidence-Based Chronic Disease Prevention in Public Health Settings Across Four Countries
Elizabeth L. Budd, Xiangji Ying, Katherine A. Stamatakis, Anna J. deRuyter, Zhaoxin Wang, Pauline Sung, Tahna Pettman, Rebecca Armstrong, Rodrigo Reis and Ross C. Brownson

110  Organizational Readiness Tools for Global Health Intervention: A Review
James W. Dearing

116  Readiness to Change Over Time: Change Commitment and Change Efficacy in a Workplace Health-Promotion Trial

ADAPTATION ASSESSMENT AND MODELS
125  Systematic, Multimethod Assessment of Adaptations Across Four Diverse Health Systems Interventions
Borsika A. Rabin, Marina McCreight, Catherine Battaglia, Roman Ayele, Robert E. Burke, Paul L. Hess, Joseph W. Frank and Russell E. Glasgow

136  The Family Check-Up 4 Health (FCU4Health): Applying Implementation Science Frameworks to the Process of Adapting an Evidence-Based Parenting Program for Prevention of Pediatric Obesity and Excess Weight Gain in Primary Care
Justin D. Smith, Cady Berkel, Jenna Rudo-Stern, Zorash Montaño, Sara M. St. George, Guillermo Prado, Anne M. Mauricio, Amanda Chiapa, Meg M. Bruening and Thomas J. Dishion

153  Using Instructional Design, Analyze, Design, Develop, Implement, and Evaluate, to Develop e-Learning Modules to Disseminate Supported Employment for Community Behavioral Health Treatment Programs in New York State
Sapana R. Patel, Paul J. Margolies, Nancy H. Covell, Cristine Lipscomb and Lisa B. Dixon

IMPLEMENTATION STRATEGIES RESEARCH
162  A Pragmatic Approach to Guide Implementation Evaluation Research: Strategy Mapping for Complex Interventions
Alexis K. Huynh, Alison B. Hamilton, Melissa M. Farmer, Bevanne Bean-Mayberry, Shannon Wiltsey Stirman, Tannaz Moin and Erin P. Finley

177  Implementation Mapping: Using Intervention Mapping to Develop Implementation Strategies
Maria E. Fernandez, Gill A. ten Hoor, Sanne van Lieshout, Serena A. Rodriguez, Rinad S. Beidas, Guy Parcel, Robert A. C. Ruiter, Christine M. Markham and Gerjo Kok
192 Implementation Climate and Time Predict Intensity of Supervision Content Related to Evidence Based Treatment
Michael D. Pullmann, Leah Lucid, Julie P. Harrison, Prerna Martin, Esther Deblinger, Katherine S. Benjamin and Shannon Dorsey

SUSTAINMENT EFFORTS
208 Using Survival Analysis to Understand Patterns of Sustainment Within a System-Driven Implementation of Multiple Evidence-Based Practices for Children’s Mental Health Services
Lauren Brookman-Frazee, Chanel Zhan, Nicole Stadnick, David Sommerfeld, Scott Roesch, Gregory A. Aarons, Debbie Innes-Gomberg, Lillian Bando and Anna S. Lau

220 Agency Leaders’ Assessments of Feasibility and Desirability of Implementation of Evidence-Based Practices in Youth-Serving Organizations Using the Stages of Implementation Completion
Lawrence A. Palinkas, Mark Campbell and Lisa Saldana
Editorial: Methods and Applications in Implementation Science

Mary E. Northridge 1*, Donna Shelley 2, Thomas G. Rundall 3 and Ross C. Brownson 4,5

1 Hansjörg Wyss Department of Plastic Surgery, NYU School of Medicine, New York University (NYU) Langone Dental Medicine - Brooklyn, Brooklyn, NY, United States, 2 Department of Population Health, NYU School of Medicine, New York, NY, United States, 3 The Center for Lean Engagement and Research (CLEAR), University of California, Berkeley School of Public Health, Berkeley, CA, United States, 4 Prevention Research Center in St. Louis, Brown School at Washington University in St. Louis, St. Louis, MO, United States, 5 Division of Public Health Sciences, Department of Surgery, Alvin J. Siteman Cancer Center, Washington University School of Medicine, St. Louis, MO, United States

Keywords: adaptation, evaluation, frameworks, implementation strategies, measurement, mechanisms, partnerships, sustainability

Editorial on the Research Topic

Methods and Applications in Implementation Science

In a classic review, Green et al. popularized the pipeline graphic that depicts the 17-year odyssey necessary for the production and transfer of knowledge from research to practice and policy (1). Still, the vetting of research through successive scientific filters does little to assure that the populations in need of evidence-based practices ever benefit from scientific advances. This Research Topic is intended to provide insights from implementation science that move beyond the clinical care of individual patients, to also take account of provider, organizational, systems, and policy levels pertaining to health and health care.

Testable theories that describe the causal pathways through which implementation strategies effect change are needed to improve the outcomes produced by evidence-based interventions (EBIs). Lewis et al. advance an innovative four-step approach to building causal pathway models that articulates the mediators, moderators, preconditions, and proximal and distal outcomes of implementation processes. Such clarity in causal pathways will allow us to understand better where, when, and why strategies have an effect on outcomes of interest.

The RE-AIM framework (2) provides important guidance for planning and assessing dimensions that influence the implementation process and potential for EBIs to impact population health. Harden et al. articulate how an updated RE-AIM framework addresses emerging implementation science priorities, such as cost and adaptation, and includes a greater focus on contextual and explanatory factors. Powell et al. present a research agenda for five priorities that need to be addressed to increase the public health impact of implementation strategies: (1) enhance methods for designing and tailoring; (2) specify and test mechanisms of change; (3) conduct more effectiveness research on discrete, multifaceted, and tailored strategies; (4) increase economic evaluations; and (5) improve tracking and reporting. For economic evaluations, the range of approaches is vast, from simple costing to full cost-effectiveness analyses. Okamura et al. report on an innovative method for calculating training and consultation costs related to delivering evidence-based treatments (EBT) that may provide insight into how systems should prioritize training efforts.

Partnerships, engagement, and collaboration (PEC) are important strategies for advancing dissemination and implementation of EBIs in clinical and community settings, but conceptual models and methods to guide design and evaluation of PECs is lacking. Huang et al. conducted a scoping review of the PEC literature that identified key domains, processes, mechanisms, and strategies for PEC, and proposed a new multilevel framework to guide future research in this area.
Mazzucca et al. assessed the research designs and methodologies used in 212 dissemination and implementation (D&I) study protocols recently published in Implementation Science. While a large majority of the protocols (77%) utilized randomized designs, and most protocols (61%) proposed quantitative and qualitative methods, only 52% reported using a theoretical framework to guide the study. Northridge et al. present a protocol for a participatory, multilevel, dynamic intervention to improve the oral health of low-income Chinese Americans, guided by two complementary, multilevel frameworks: Consolidated Framework for Implementation Research (CFIR) (3) and Implementation Outcomes Framework (IOF) (4). Lee et al. utilized a novel multiphase, explanatory sequential mixed methods design to provide deeper understanding of how complex multisector partnerships impact population health outcomes in an evaluation of the Massachusetts Prevention and Wellness Trust Fund.

As per the public health adage, “what gets measured gets done,” (5) progress in implementation requires the development of practical measures that are both reliable and valid. Budd et al. developed and tested a tool for measuring the contextual factors related to evidence-based practice across four countries (Australia, Brazil, China, and the United States), and found variability in reliability across domain and country, suggesting that some items are highly generalizable, while others are less so. Dearing conducted a review of 30 available organizational readiness tools, noting that even as most measure capacity, few measure organizational motivation. Helfrich et al. assessed organizational readiness to change over two waves in a workplace health promotion trial, and found that change commitment declined significantly at both intervention and control sites over time, even as wellness-program effort increased significantly at intervention sites.

Adapting EBIs to the local context is a necessary step to facilitate adoption and implementation. Approaches are needed that promote a systematic approach to documenting and evaluating the adaptation process. Rabin et al. make an important contribution by describing a multilevel, multimethod adaptation approach across four health systems, guided by the Stirman framework (6) for adaptation and modification and expanded using concepts from the RE-AIM framework (2). The modified adaptation model showed promise in capturing adaptation across a range of projects and content areas. To scale-up an evidence-based parenting program for prevention of pediatric obesity, Smith et al. report on the multiyear process of adaptation to a new clinical target and service delivery system. In a study of behavioral health treatment, Patel et al. apply an instructional design framework in the development and evaluation of e-learning modules as either a single component or one strategy in a multifaceted approach for training in evidence-based practices (EBPs).

Detailed specification of implementation strategies is a challenge, especially for complex, multilevel interventions that use multiple strategies. Huynh et al. describe a five-step method for mapping intervention strategies and demonstrate its use with a study of the implementation of a cardiovascular toolkit. Fernandez et al. introduce Implementation Mapping, which provides a systematic process for developing strategies to improve the adoption, implementation, and maintenance of evidence-based interventions in real-world settings. Pullmann et al. report on findings from a study of the impact of clinical supervision to improve the adoption of EBT for child mental health problems. Findings point to the importance of a supportive organizational climate in predicting supervisory EBT intensity.

Brookman-Frazee et al. contribute to the limited research on EBP sustainment in mental health services long after implementation, illustrating a novel application of survival analysis to administrative claims data in system-driven implementation of multiple EBPs. Finally, Palinkas et al. point to opportunities for using agency leader models to develop strategies to facilitate implementation of evidence-based and innovative practices for children and adolescents, guided by the Stages of Implementation Completion framework (7). Our hope is that this collection advances the field.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Northridge, Shelley, Randall and Brownson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
From Classification to Causality: Advancing Understanding of Mechanisms of Change in Implementation Science

Cara C. Lewis1,2,3*, Predrag Klasnja1†, Byron J. Powell1, Aaron R. Lyon3, Leah Tuzzio1, Salene Jones5, Callie Walsh-Bailey1 and Bryan Weiner6

1 Kaiser Permanente Washington Health Research Institute, Seattle, WA, United States, 2 Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, United States, 3 Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States, 4 Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, 5 Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States, 6 Department of Global Health, University of Washington, Seattle, WA, United States

Background: The science of implementation has offered little toward understanding how different implementation strategies work. To improve outcomes of implementation efforts, the field needs precise, testable theories that describe the causal pathways through which implementation strategies function. In this perspective piece, we describe a four-step approach to developing causal pathway models for implementation strategies.

Building causal models: First, it is important to ensure that implementation strategies are appropriately specified. Some strategies in published compilations are well defined but may not be specified in terms of its core component that can have a reliable and measureable impact. Second, linkages between strategies and mechanisms need to be generated. Existing compilations do not offer mechanisms by which strategies act, or the processes or events through which an implementation strategy operates to affect desired implementation outcomes. Third, it is critical to identify proximal and distal outcomes the strategy is theorized to impact, with the former being direct, measurable products of the strategy and the latter being one of eight implementation outcomes (1). Finally, articulating effect modifiers, like preconditions and moderators, allow for an understanding of where, when, and why strategies have an effect on outcomes of interest.

Future directions: We argue for greater precision in use of terms for factors implicated in implementation processes; development of guidelines for selecting research design and study plans that account for practical constructs and allow for the study of mechanisms; psychometrically strong and pragmatic measures of mechanisms; and more robust curation of evidence for knowledge transfer and use.

Keywords: implementation, mechanism, mediator, moderator, theory, causal pathway, strategy

In recent years, there has been growing recognition of the importance of implementing evidence-based practices as a way to improve the quality of health care and public health. However, the results of implementation efforts have been mixed. About two-thirds of efforts fail to achieve the intended change (2), and nearly half have no effect on outcomes of interest (3). Implementation strategies are
often mismatched to barriers [e.g., training, a strategy that could affect implementation outcomes through changes in an individual's knowledge (intrapersonal-level), is used inappropriately to address an organizational-level barrier like poor culture] (4), and implementation efforts are increasingly complex and costly without enhanced impact (5). These suboptimal outcomes are due, in large part, to the dearth of tested theory in the field of implementation science (6). In particular, the field has a limited understanding of how different implementation strategies work—the specific causal mechanisms through which implementation strategies influence care delivery [7; Lewis et al. (under review)]3. As a consequence, implementation science has been limited in its ability to effectively inform implementation practice by providing guidance about when and in what contexts specific implementation strategies should be used and, just as importantly, when they should not.

The National Academy of Science defines “science” as “the use of evidence to construct testable explanations and predictions of natural phenomena, as well as the knowledge generated through this process.” (8) The field of implementation has spent the past two decades building and organizing knowledge, but we are far from having testable explanations that afford us the ability to generate predictions. To improve outcomes of implementation efforts, the field needs testable theories that describe the causal pathways through which implementation strategies function (6, 9). Unlike frameworks, which offer a basic conceptual structure underlying a system or concept (10), theories provide a testable way of explaining phenomena by specifying relations among variables, thus enabling prediction of outcomes (10, 11).

Causal pathway models represent interrelations among variables and outcomes of interest in a given context (i.e., the building blocks of implementation theory). Specifying the structure of causal relations enables scientists to empirically test whether the implementation strategies are operating via theorized mechanisms, how contextual factors moderate the causal processes through which implementation strategies operate, and how much variance in outcomes is accounted for by those mechanisms. Findings from studies based on causal models can, over time, both help the field develop more robust theories about implementation processes and advance the practice of implementation by addressing key issues. For instance, causal models can do the following: (1) inform the development of improved implementation strategies, (2) identify mutable targets for new strategies, (3) increase the impact of existing strategies, and (4) prioritize which strategies to use in which contexts.

In this perspective piece, we propose an approach to theory development by specifying, in the form of causal pathway models, hypotheses about the causal operation of different implementation strategies in various settings, so that these hypotheses can be tested and refined. Specifically, we offer a four-step process to developing causal pathway models for implementation strategies. Toward this end, we argue the field must move beyond having lists of variables that can rightly be considered determinants [i.e., factors that obstruct or enable change in provider behavior or health-care delivery processes (12)], and toward precise articulation of mediators, moderators, preconditions, and (proximal versus distal) outcomes (see Table 1 for definitions).

### BUILDING CAUSAL PATHWAY MODELS

Our perspective draws upon Agile Science (13, 14)—a new method for developing and studying behavioral interventions that focuses on intervention modularity, causal modeling, and efficient evaluations to generate empirical evidence with clear boundary conditions (in terms of population, context, behavior, etc.) to maximize knowledge accumulation and repurposing. Agile Science has been used to investigate goal-setting interventions for physical activity, engagement strategies for mobile health applications, depression interventions for primary care, and automated dietary cues to promote weight loss (13, 15). Applied to implementation strategies, Agile Science-informed causal pathway diagram modeling consists of at least four steps: (1) specifying implementation strategies; (2) generating strategy-mechanism linkages; (3) identifying proximal and distal outcomes; and (4) articulating moderators and preconditions. To demonstrate this approach, we offer examples of causal pathway models for a set of three diverse implementation strategies (see Figure 1). The strategies are drawn from the following example. A community mental health center is planning to implement measurement-based care in which providers solicit patient-reported outcome data [e.g., Patient Health Questionnaire 9-item depression symptom severity measure (16)] prior to clinical encounters to inform treatment (17). The community mental health center plans to use training, financial penalty (disincentives), and audit and feedback as they are common strategies used to support measurement-based care implementation (18).

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanism</td>
<td>Process or event through which an implementation strategy operates to affect desired implementation outcomes</td>
</tr>
<tr>
<td>Precondition</td>
<td>Factor that is necessary in order for an implementation mechanism to be activated</td>
</tr>
<tr>
<td>Determinant</td>
<td>Also commonly referred to as “barriers” and “facilitators,” a factor that enables or hinders the implementation strategy from eliciting the desired effect</td>
</tr>
<tr>
<td>Mediator</td>
<td>Intervening variable that may account for the relationship between the implementation strategy and the implementation outcome</td>
</tr>
<tr>
<td>Moderator</td>
<td>Factor that increase or decrease the level of influence of an implementation strategy</td>
</tr>
<tr>
<td>Proximal outcome</td>
<td>The product of the implementation strategy that is realized because of its specific mechanism of action, the most immediate, observable outcome in the causal pathway</td>
</tr>
<tr>
<td>Distal outcome</td>
<td>Outcomes that the implementation processes is ultimately intended to achieve, not the most immediate outcome in the causal pathway</td>
</tr>
</tbody>
</table>

### Step 1: Specifying Implementation Strategies

The Expert Recommendations for Implementing Change study yielded a compilation of 73 implementation strategies (19)
developed by a multidisciplinary team through a structured literature review (20), Delphi process, and concept mapping exercise (19, 21, 22). Thus, there exists a solid foundation of strategies that are conceptually clear and well defined. However, the compilation was never explicitly linked to mechanisms. Following Kazdin (7), we define “mechanisms” as the processes or events through which an implementation strategy operates to effect desired implementation outcomes. Upon careful examination, it seems many strategies are not well enough specified to be linked to mechanisms in a coherent manner, a key step in causal model building.
For instance, the compilation of 73 strategies lists “learning collaboratives,” a general approach for which the discrete strategies or core components are underspecified. This makes it difficult to identify their precise mechanisms of action (23). Underspecified strategies also leave the field vulnerable to inappropriately synthesizing data across studies (24, 25).

In our case example, training is a strategy that is underspecified. We adapted procedures from Michie et al. (26) to guide strategy specification recommending that each strategy be assessed for whether it: (1) aims to promote the adoption, implementation, sustainment, or scale-up of an evidence-based practice; (2) is a proposed “active ingredient” of adoption, implementation, sustainment, or scale-up; (3) represents the smallest component while retaining the proposed active ingredient; (4) can be used alone or in combination with other discrete strategies; (5) is observable and replicable; and (6) can have a measurable impact on specified mechanisms of implementation (and, if so, whether putative mechanisms can be listed). If strategies do not meet these criteria, they require revision and further specification. This could involve suggesting alternative definitions, eliminating an implementation strategy altogether, or articulating a new, narrower strategy that is a component or a type of the original strategy. Training would meet all but the third and sixth criteria (listed previously), because training can be comprised of several active ingredients (e.g., didactics, modeling, role play/rehearsal, feedback, shadowing) each of which may operate on an unique mechanism. In this case, training ought to be more narrowly defined to make clear its core components.

**Step 2: Generating Strategy-Mechanism Linkages**

Once specified, an implementation strategy needs to be linked to the mechanisms hypothesized to underlie its functioning. Mechanisms explain how an implementation strategy has an effect by describing the actions that lead from the administration of the strategy to the implementation outcomes (see Table 1 for definitions). Statistically speaking, mechanisms are always mediators, but mediators may not be mechanisms. Similarly, moderators can point toward mechanisms but are not themselves reliably mechanisms. Determinants may explain why an implementation strategy did or did not have an effect, but mechanisms explain how a strategy had an effect, by, for example, altering the status of a determinant. Determinants are naturally occurring, and often but not always, malleable factors that could prevent or enable the strategy to affect the desired outcomes. Mechanisms are intentionally activated by the application of an implementation strategy and can operate at different levels of analysis, such as at the levels of intrapersonal (e.g., learning), interpersonal (e.g., sharing), organizational (e.g., leading), community (e.g., restructuring), and macro policy (e.g., guiding) (27).

For an implementation effort to be successful, chosen strategies should be compatible with and able to act on the local determinants [e.g., provider habit (determinant) is addressed with clinical decision support (strategy) via self-reflection/reflecting (mechanism)]. Although commonly used in implementation science, we propose that the notion of a determinant is insufficiently specific as researchers have used it to refer to at least two types of variables in a causal process: proximal outcomes and effect modifiers (see text footnote 1). Our discussion below uses these more precise terms instead.

Most implementation strategies likely act via multiple mechanisms, although it remains an empirical question whether one mechanism is primary and others are ancillary. It is also likely that the same mechanism might be involved in the operation of multiple implementation strategies. Initial assessment of strategy-mechanism linkages is made in the context of the broader scientific knowledge base about how a strategy produces an outcome (7). For instance, many strategies have their own literature base (e.g., audit and feedback) (28) that offer theoretical and empirical insights about which mechanisms might be underlying the functioning of those strategies [e.g., reflecting, learning, and engaging (28)]. Effort should always be made to draw upon and test existing theories, but if none offer sufficient guidance, hypothesizing variables that may have causal influence remains critical. In this way, over time, the initially formulated strategy-mechanism linkages can be reassessed and refined as studies begin to test them empirically. While such empirical evaluations are currently rare—across two systematic reviews of implementation mechanisms, only 31 studies were identified and no mechanisms were empirically established (see text footnote 1)—the causal pathway models we propose here are explicitly intended to facilitate evaluations of the mechanistic processes through which implementation strategies operate.

**Step 3: Identifying Proximal and Distal Outcomes**

Implementation scientists have isolated eight outcomes as the desired endpoints of implementation efforts: acceptability, feasibility, appropriateness, adoption, penetration, fidelity, cost, and sustainability (1). Many of these outcomes are appropriately construed as latent variables, but others are manifest/observable in nature (30); a recent systematic review offers measures of these outcomes and measure meta-data (31). In terms of the causal processes through which implementation strategies operate, these outcomes are often best conceptualized as distal outcomes that the implementation process is intended to achieve, and each of them may be more salient at one phase of implementation than another. For instance, with the Exploration, Preparation, Implementation, Sustainment Framework (32), acceptability of an evidence-based practice may be most salient in the exploration phase, whereas fidelity may be the goal of an implementation phase. Despite the plausible temporal interrelations among the outcomes, mounting evidence indicates that not all implementation strategies influence each of the aforementioned outcomes (e.g., workshop training can influence adoption but not fidelity) (33). To fully establish the plausibility of an implementation mechanism and a testable causal pathway, proximal outcomes must be expounded.

Proximal outcomes are direct, measurable, and typically observable, products of the implementation strategy that occur because of its specific mechanism of action. That is, affecting a proximal outcome in the intended direction can confirm/disconfirm activation of the putative mechanism, offering a low-inference way to establish evidence for a theorized mechanism. Most often, mechanisms themselves cannot be directly measured, forcing (either high-inference assessment or) reliance on the observation of change in a proximal outcome of interest. For instance, didactic education, as an active ingredient of training, acts primarily...
through the mechanism of learning on the proximal outcome of knowledge to influence the distal implementation outcome of perceived acceptability or even adoption. Practice with feedback acts through the mechanism of reflecting on proximal outcomes of skills and confidence to influence the distal implementation outcome of adoption or even fidelity. To identify proximal outcomes, one must answer the question, “How will I know if this implementation strategy had an effect via the mechanism that I think it is activating?” or “What will be different if the hypothesized mechanisms for this strategy is at play?” It is very common for mechanisms and proximal outcomes to be conflated in the literature given that researchers often test mediation models examining the impact of a strategy on a distal implementation outcome via a more proximal outcome. The way we are using the terms, a mechanism is a process through which an implementation strategy operates, and a proximal outcome is a measurable effect of that process that is in the causal pathway toward the distal implementation outcomes.

Step 4: Articulating Effect Modifiers

Finally, there are two types of effect modifiers that are important to articulate, both of which can occur across multiple levels of analysis: moderators and preconditions. Moderators are factors that increase or decrease the level of influence of an implementation strategy on an outcome. See Figure 1 in which an example for intra-individual and organizational-level moderators for audit and feedback are articulated. Theoretically, moderators are factors that interact with a strategy’s mechanism of action, even if exactly how they interact mechanistically are not understood. Preconditions are factors that are necessary for an implementation mechanism to be activated at all (see Figure 1). They are necessary conditions that need to be in place for the causal process that leads from an implementation strategy to its proximal and distal outcomes to take place. Both moderators and preconditions are most often mischaracterized as “determinants” in the implementation science literature base, which may limit our ability to understand the nature of the relations between a strategy and the individual and contextual factors that modify its effects, and, in turn, where, when, and why strategies have an effect on outcomes of interest.

FUTURE DIRECTIONS: WHAT THE FIELD OF IMPLEMENTATION NEEDS TO FULLY ESTABLISH ITSELF AS A SCIENCE

In order to fully establish itself as a science by offering testable explanations and enabling the generation of predictions, we offer four critical steps for the field of implementation: (1) specify implementation strategies; (2) generate implementation strategy-mechanism linkages; (3) identify proximal and distal outcomes; and (4) articulate effect modifiers. In addition to these steps, we suggest that future research should strive for the generation of precise terms for factors implicated in implementation processes and use them consistently across studies. In a systematic review of implementation mechanisms, researchers conflated preconditions, predictors, moderators, mediators, and proximal outcomes (see text footnote 1). In addition, there is room for the field to develop guidelines for selecting research designs and study plans that account for practical constraints of the contexts in which implementation is studied and allow for mechanism evaluation. The types of causal pathway models that we advocated for here, paired with an understanding of the constraints of a particular study site, would enable researchers to select appropriate methods and designs to evaluate hypothesized relations by carefully considering the temporal dynamics such as how often a mechanism should be measured and how much the outcome is expected to change and when.

In order to truly advance the field, much work needs to be done to identify or develop psychometrically strong and pragmatic measures of implementation mechanisms. Empirically evaluating causal pathway models requires psychometrically strong measures of mechanisms that are also pragmatic, yet none of the seven published reviews of implementation-relevant measures focus on mechanisms. It is likely that measure development will be necessary to advance the field. Finally, implementation science could benefit from the building of more robust curation of evidence for knowledge transfer and use. Other fields house web-based databases for collecting, organizing, and synthesizing empirical findings [e.g., Science of Behavior Change (34)].

In doing so, fields can accumulate knowledge more rapidly and users of knowledge can determine what is working, when, and why, as well as what generalizes and what does not. Such curation of evidence can more efficiently lead to the development of improved implementation strategies (e.g., through strategy specification), identification of mutable targets for new strategies (e.g., mechanisms revealed for existing strategies that may not be pragmatic), and prioritization of strategy use for a given context (e.g., given knowledge of preconditions and moderators).

AUTHOR CONTRIBUTIONS

CL and PK are co-first authors, who co-led manuscript development. CL and BW are co-PIs on an R01 proposal that led to the inception of this manuscript. All authors (CL, PK, BP, AL, LT, SJ, CW-B, and BW) contributed to idea development, writing, and editing of this manuscript and agreed with its content.

ACKNOWLEDGMENT

BP would like to acknowledge funding from the National Institute of Mental Health (K01MH113806).

REFERENCES


34. Columbia University Medical Center, National Institutes of Health. *Science of Behavior Change [Internet].* Columbia University Medical Center [cited 2017 Nov 6]. Available from: https://scienceofbehaviorchange.org/

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer TW declared a past co-authorship with one of the author BP to the handling Editor.

Copyright © 2018 Lewis, Klasnja, Powell, Lyon, Tuzzio, Jones, Walsh-Bailey and Weiner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
RE-AIM in Clinical, Community, and Corporate Settings: Perspectives, Strategies, and Recommendations to Enhance Public Health Impact

Samantha M. Harden1*, Matthew Lee Smith2,3,4, Marcia G. Ory2,3, Renae L. Smith-Ray5, Paul A. Estabrooks6 and Russell E. Glasgow7

1 Physical Activity Research and Community Implementation, Human Nutrition, Foods, and Exercise, Virginia Tech, Blacksburg, VA, United States, 2 Center for Population Health and Management, Texas A&M University, College Station, TX, United States, 3 Department of Environmental and Occupational Health, School of Public Health, Texas A&M University, College Station, TX, United States, 4 Department of Health Promotion and Behavior, College of Public Health, The University of Georgia, Athens, GA, United States, 5 Walgreens Center for Health and Wellbeing Research, Deerfield, IL, United States, 6 Department of Health Promotion, College of Public Health, University of Nebraska Medical Center, Omaha, NE, United States, 7 Department of Family Medicine, School of Medicine, University of Colorado, Aurora, IL, United States

The RE-AIM Framework is a planning and evaluation model that has been used in a variety of settings to address various programmatic, environmental, and policy innovations for improving population health. In addition to the broad application and diverse use of the framework, there are lessons learned and recommendations for the future use of the framework across clinical, community, and corporate settings. The purposes of this article are to: (A) provide a brief overview of the RE-AIM Framework and its pragmatic use for planning and evaluation; (B) offer recommendations to facilitate the application of RE-AIM in clinical, community, and corporate settings; and (C) share perspectives and lessons learned about employing RE-AIM dimensions in the planning, implementation, and evaluation phases within these different settings. In this article, we demonstrate how the RE-AIM concepts and elements within each dimension can be applied by researchers and practitioners in diverse settings, among diverse populations and for diverse health topics.

Keywords: translation, health promotion, knowledge transfer, implementation science, evaluation framework, dissemination and implementation research

INTRODUCTION

Dissemination and implementation (D&I) research addresses the “how and why” related to strategies for information sharing (dissemination) and intervention integration (implementation) for the purposes of enhancing evidence-based program delivery and population health (1–5). The advancement of D&I science requires a focus on the wide-scale adoption, implementation, and generalizability of program and policy impacts. With well over 100 different models and frameworks utilized in the field (6), researchers and practitioners can become overwhelmed when selecting (and attempting to apply) the most appropriate model/framework for their scientific inquiry or initiative.1

The purposes of this article are to: (A) provide a brief overview of the RE-AIM Framework and its pragmatic use for planning and evaluation; (B) offer recommendations to facilitate the application of

1 www.dissemination-implementation.org.
RE-AIM in clinical, community, and corporate settings; and (C) share perspectives and lessons learned about employing RE-AIM elements in the planning, implementation, and evaluation phases within these different settings. In this article, we demonstrate how RE-AIM concepts and elements can be applied by researchers and practitioners in diverse settings, among diverse populations, and for diverse health topics.

THE RE-AIM FRAMEWORK

The RE-AIM Framework (7, 8) is often used in D&I research (9, 10), which encompasses essential translational research elements. RE-AIM was identified as the most frequently used model or framework between 2000 and 2016 for D&I grant applications submitted to the National Institutes of Health and Centers for Disease Control and Prevention (CDC) (11). This widespread use is, in part, due to the flexibility to address different public health concerns in a practical manner understandable by practitioners and policy makers. The acronym RE-AIM stands for reach (How do I reach those who need a specific intervention?), efficacy/effectiveness (How do I know my intervention is working?), adoption (How do I design for dissemination and develop organizational support to deliver my intervention?), implementation (How do I ensure the intervention is feasible and delivered properly?), and maintenance (How do I ensure long-term benefits and institutionalization of the intervention and continued community capacity for D&I?).

Applying RE-AIM challenges researchers and practitioners to ask fundamental questions about complex issues before, during, and after the implementation of a putative program in “real world” settings. Among the many strengths of RE-AIM is its robust structure that facilitates broad use across settings (e.g., organization, regional, rural), populations (e.g., age, race/ethnicity, occupation/role), topics (e.g., disease, behavior), and interventions (e.g., demonstration, experimental, translational, longitudinal, multi-level). While the basic RE-AIM dimensions have remained constant since its development in the 1990s (7), its use has evolved over time with new applications in clinical (12), community (13), and corporate (14) settings. A recent systematic review (15) reported health-care (49%) and community (46%) settings applied RE-AIM in empirical or evaluative interventions most frequently; however, no such interventions were reported in corporate settings. As such, efforts are needed to understand the use of RE-AIM in multiple settings. Researchers and practitioners are encouraged to use the RE-AIM framework for beginning with the end in mind, designing for dissemination, and evaluating relevant dimensions across intervention and setting factors. Such deliberate RE-AIM application will contribute to the replicability and generalizability of planned interventions and thus yield optimal public health impact.

PRAGMATIC USE OF RE-AIM FOR PLANNING AND EVALUATION

The RE-AIM Framework can be used to direct the planning of new or ongoing interventions and systematic evaluations that include a complex interplay of individual and organizational outcomes (10). Fully employing RE-AIM can speed the translation of effective interventions in practice settings, while demonstrating impact and representativeness (9, 10). Yet, utilizing the full framework may require substantial human, data, and analytic resources that may not be available or feasibly acquired across typical clinical, community, or corporate settings (16). This is especially true in settings where decision-making may be based on a small subset of RE-AIM dimensions coupled with organizational priorities and resources.

Settings must consider the temporality of assessment for each RE-AIM dimension, which may need to occur prospectively, concurrently, and/or retrospectively to determine the impact of an initiative. While employing RE-AIM before an intervention begins is ideal to ensure careful and strategic local planning, in some cases this is not possible. Some organizational practices may be the result of opportunistic intervention, rollout from a central administrative site, innovation testing; corporate, policy, or organizational directive; or quality control and enhancement—each of which has distinct challenges in aligning the evaluation with initiative strategies.

Figure 1 illustrates the application of RE-AIM based on the starting temporal stage of an intervention or initiative, that is if the RE-AIM planning and evaluation is initiated before, during, or after an initiative has been completed. Each temporal starting point includes reflective processes in which researchers or practitioners can gather information (assess) and think critically about the relevance of each RE-AIM dimension (plan). Each stage also includes active processes where those applying RE-AIM can initiate and implement plans for interventions or initiatives (do), process gathered information based on predetermined criteria (evaluate), and engage partners and stakeholders in interpretation to support decision-making (report). The bidirectional arrow along the temporal stages indicate the iterative nature of these processes, each building upon one another to provide cumulative input for advancement and refinement based evolving priorities, challenges, and observed impacts (2, 17). The importance of Figure 1 is to address the iterative nature of applying RE-AIM in planning and evaluation and how new data are taken into consideration and used to engage in a planning and action process.

As the bidirectional arrow suggests, the end of an initiative is the beginning of another (i.e., sustained implementation, adapted implementation, or implementation of an alternative solution), thus the process is cyclical and ongoing. While it is not feasible to always employ RE-AIM before an intervention or initiative begins, this figure indicates that the process can begin at any temporal stage. At all stages, researchers and practitioners are encouraged to APDER: Assess (using relevant RE-AIM dimensions and available data); Plan (based on best science, program priorities, stakeholder and organizational values, and available resources); Do (based on predetermined plans using defined procedures/protocols and supporting appropriate adaptations as needed during implementation); Evaluate (based on criteria necessary for decision-making and iterative adjustment); and Report (to, and plan for follow-up with, key stakeholders).
The RE-AIM website\(^2\) hosts a planning and evaluation document, which includes prompts and considerations across all five RE-AIM dimensions by temporal stage within a project,\(^3\) which is also available as a supplemental table to this manuscript (see Appendix A in Supplementary Material). Selected examples of common pragmatic considerations are described below.

Engaging key stakeholders (e.g., policy makers, service delivery personnel, members from the population intended to benefit from the work) is important for guiding pragmatic evaluations using RE-AIM. Researchers and practitioners should partner with organizational decision-makers to identify the necessary information required to determine priorities, justify the need for intervention, sustain implementation, and/or broaden adoption. For example, if a strategy is delivered by a single organization with a centralized delivery infrastructure, issues related to reach and effectiveness (as well as implementation costs and sustainability) may be more relevant than adoption \(^{18}\). Conversely, when attempting to scale-up or scale-out an effective intervention across a number of sites (within or across organizations), issues related to implementation quality/fidelity and adoption may be considered more important than documenting the intervention’s effectiveness in new and diverse settings \(^{19, 20}\).

Pragmatically measuring RE-AIM outcomes \(^{21}\) includes leveraging data already collected within the organizational setting to reduce evaluation costs and enhance local relevance. For example, imagine a health-care system will employ a multi-leveled intervention to enhance diabetes control by promoting physical activity. The intervention includes screening, brief counseling, referral to internal or external resources for physical activity. A pragmatic evaluation of this approach may include using electronic health records to assess the reach and representativeness of participants, changes in physical activity based on clinical screenings over time, and the number of referrals made \(^{22}\). Based on priorities and available resources, it may be less pragmatic for the health-care system to assess patients’ use of external resources for physical activity or their actual physical activity levels. However, if a similar multi-level intervention were implemented in a community setting, accessing electronic health records may be politically, legally or cost-prohibitive, or less relevant; rather, documenting participants’ physical activity with pedometers/accelerometers and tracking facility utilization are prioritized.

Available resources for evaluation are often limited in “real world” non-academic community and clinical settings. In most settings, resources are allocated to the intervention’s delivery and management to maximize enrollment/engagement. Therefore, the pragmatic selection and use of existing measures is helpful to reduce data collection burden. However, the use of existing measures can also introduce resource needs associated with data extraction, case de-identification, and statistical analyses and data management that may exceed organizational skillsets and typical reporting procedures.

EXAMPLES OF RE-AIM IN DIFFERENT SETTINGS

In this section, we provide examples of RE-AIM application in three major types of settings. In addition to these examples, Table \(1\) contains additional recommendations for using RE-AIM by temporal stages of an intervention (i.e., before, during, after) across clinical, community, and corporate settings. The purpose of this table is to document the consistency of topics to be considered when applying RE-AIM across settings, while highlighting

\(^1\)www.re-aim.org.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Examples of applying RE-AIM dimension(s) in different settings across different phases of projects.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project stage</strong></td>
<td><strong>Clinical</strong></td>
</tr>
<tr>
<td><strong>BEFORE IMPLEMENTATION</strong></td>
<td></td>
</tr>
<tr>
<td>Consider the project impact on all RE-AIM dimensions and prioritize the focus for planning and evaluation</td>
<td>Example: stakeholders’ interest in intervention reach and representativeness within the setting</td>
</tr>
<tr>
<td></td>
<td>Considerations: may need to conduct sensitivity analyses to determine sample size because of issues like inconsistent coding. There may be coding inconsistencies that influence the numerator or denominator, and all data may not be available for the desired study.</td>
</tr>
<tr>
<td></td>
<td>Prioritization: although reach is an important dimension to consider, in this example, the team prioritizes effect of the behavioral outcome</td>
</tr>
<tr>
<td>Determine how each dimension will be included in the project: describe, assess, and/or intervene</td>
<td>Example: decision made to intervene to improve adoption, describe effect, and assess implementation fidelity</td>
</tr>
<tr>
<td></td>
<td>Intervene: health-care organization is implementing new protocol for nursing rounds. Some clinics receive additional intervention to improve adoption of the protocol.</td>
</tr>
<tr>
<td></td>
<td>Describe or measure the effect of the new rounding protocol (i.e., did it achieve outcome of interest).</td>
</tr>
<tr>
<td>Develop data collection and reporting procedures and timelines for selected RE-AIM dimensions</td>
<td>Consider the metrics of interest and how data will be transferred.</td>
</tr>
<tr>
<td></td>
<td>Consider if HIPAA compliance or BAA/DUA* are needed.</td>
</tr>
<tr>
<td></td>
<td>Determine the appropriate timeline for observing outcomes (e.g., a full year of observation may be needed to see change in clinical outcomes).</td>
</tr>
<tr>
<td>Engage all project staff and partners in processes to ensure transparency, equity, compliance with regulations, and support (ongoing throughout the project)</td>
<td>Example: determine appropriate stakeholders and where, when, how, and why they will be engaged</td>
</tr>
<tr>
<td></td>
<td>Consider that perhaps it is not appropriate to engage patients with an electronic medical record update.</td>
</tr>
</tbody>
</table>

(Continued)
### TABLE 1 | Continued

<table>
<thead>
<tr>
<th>Project stage</th>
<th>Clinical</th>
<th>Community</th>
<th>Corporate</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan for sustainability and generalizability from the outset</td>
<td>Consider how intervention- and assessment- components can be implemented in settings with different histories, resources, workflows. Plan to communicate results with stakeholders providing guidance and align reporting of information with data needed for decision-making for sustainability.</td>
<td>Develop a coalition or advisory board to be engaged throughout the process, including those not directly involved in the project, to identify information and resources needed to increase the likelihood of sustainability.</td>
<td>Include staff with clinical expertise to be engaged throughout the process, including those not directly involved in the project.</td>
<td>Design for feasibility, success, and dissemination that addresses each of RE-AIM dimensions. Design the intervention to be broadly applied within and across settings.</td>
</tr>
</tbody>
</table>

### DURING IMPLEMENTATION/ITERATIVE ASSESSMENT AND ADJUSTMENT

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitor data periodically and at key points for each dimension (emphasis on priority dimensions)</td>
<td>Have brief (perhaps “automated”), ongoing data collection. Use rapid, pragmatic assessments to identify reasons for initial results.</td>
<td>Conduct training for program delivery staff about data collection procedures including data completion and quality checks. Routinely export available data from administrative records and secondary sources to track real-time changes.</td>
<td>Have brief “automated” ongoing data collection from routine company records. When supplementary outcome measures are used, conduct training for program delivery staff about data collection procedures including data completion and quality checks. Routinely export available data from administrative records and secondary sources to track real-time changes.</td>
<td>Pragmatic, timely, and low-resource data collection for ongoing decision-making and engagement in the PDSA cycle over time and dimensions.</td>
</tr>
<tr>
<td>Track implementation and costs as well as fidelity to core components if those are priority dimensions</td>
<td>Discuss and implement low burden cost assessments (interviews, tracking, observations) at key time points.</td>
<td>Develop systems for fidelity monitoring (observation) and adherence to delivery protocol. Programs that breach fidelity are subject to additional unplanned costs (e.g., cost per participant increases if workshops are not filled to capacity).</td>
<td>Track implementation and variability across sites. Routinely compare outcomes across a random sample of sites as a way of identifying unanticipated fluctuations and potential protocol deviations.</td>
<td>Real-time issues can be addressed more rapidly. Avoids type 3 error (concluding that intervention did not work when perhaps delivery was not consistent with evidence-based components).</td>
</tr>
<tr>
<td>Perform ongoing assessments of project evolution and adaptations</td>
<td>Probe adaptations to address each RE-AIM dimension. Track implementation and impact over time and across settings and staff.</td>
<td>Routinely export available data from administrative records and secondary sources to track real-time progress. Regularly debrief with program deliverers and organizational partners to identify (and adapt to address) unforeseen challenges.</td>
<td>Track implementation and impact over time and across settings and staff. Collect stories and “positive deviance” examples to inspire other settings.</td>
<td>Need to capture real-world adaptations to systematically collect data on how, why, when, and by whom changes are being implemented in the field.</td>
</tr>
<tr>
<td>Reconsider the intervention impact on (and priorities for) all RE-AIM dimensions</td>
<td>Use both quantitative and qualitative assessments. In applied cases, use “good enough” methods—ballpark estimates make them work when “gold standard” methods are not feasible.</td>
<td>Assess whether the number of participants reached will enable meaningful outcomes to be observed and adjust recruitment/delivery accordingly. Discuss project progress with program deliverers, partnering organizations, and other key stakeholders regularly to ensure transparency and identify changes in priorities for the project.</td>
<td>Assess program impact on “bottom line” and estimated return-on-investment. Discuss project progress with program deliverers, different locations, and other key stakeholders regularly to ensure transparency and identify changes in priorities for the project.</td>
<td>Continued discussion with stakeholders ensures that the appropriate impact is being achieved. Ongoing considerations of which dimension to intervene, describe, or assess, particularly for long-term intervention work.</td>
</tr>
<tr>
<td>Decide if adaptations are needed to address problems with outcomes on one or more RE-AIM dimensions</td>
<td>Pilot and then implement intervention or implementation strategy adaptations needed to improve performance, and track their impact.</td>
<td>Assess the appropriateness of participants engaged in the intervention to determine if appropriate and equitable outcomes are observed. Depending on what is seen, there may be implications for refining participant recruitment and retention procedures.</td>
<td>Test different intervention or implementation strategy adaptations needed to improve performance, and track their impact. Track innovations.</td>
<td>Prioritize adaptations and test their impact across dimensions (see Figure 1).</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AFTER IMPLEMENTATION/SUMMATIVE</td>
<td>Evaluate the impact on all relevant RE-AIM dimensions. Consider subgroup as well as overall effects. Consider overall impact on quality of life and patient-centered outcomes. Include balancing measures.</td>
<td>Begin with priority dimensions and “low-hanging fruit”. Reach and implementation measures may be easily assessed, whereas adoption and maintenance may require more in-depth processes to identify.</td>
<td>Consider subgroup effects in addition to overall outcomes. Based on findings, target intervention to streamline resources and impact.</td>
<td>Return to RE-AIM plan and summarize accordingly.</td>
</tr>
</tbody>
</table>

(Continued)
the unique factors framing the contextualization of RE-AIM within settings.

Clinical Health-care Setting

Esteemed professional organizations and societies (e.g., The Institute of Medicine, National Academies of Medicine, and Society of Behavioral Medicine) have called for health systems to assess key health behaviors, mental health, and social measures, and address an actionable set of social determinants of health. Leveraging these opportunities, the My Own Health Report (MOHR) consortium tested a brief, evidence-based online and interactive health risk assessment and feedback tool (MyOwnHealthReport.org). The online aid included patient-reported items on health risk behaviors, mental health, substance use, demographics, and patient preferences (23).

The MOHR project tested the interactive patient-report and feedback system in a cluster randomized trial of 18 primary care clinics across five states. RE-AIM was used to plan, adapt, and evaluate the system using a low-cost pragmatic implementation strategy. RE-AIM was used in the planning stages to develop strategies feasible for low-resource settings with patients most in need (e.g., federally qualified health centers and other diverse clinics including rural, suburban, and urban clinics). Inclusion criteria were purposely broad for clinics and patients, and time demands on patients and staff were kept to a minimum. The implementation plan involved a high degree of flexibility and allowed each clinic to recruit patients, administer the MOHR, simultaneously provide feedback, use assessment/feedback modalities, select languages (English or Spanish), and place in their clinic workflow. In terms of RE-AIM, this plan addressed reach, adoption, and implementation issues.

RE-AIM was used iteratively to monitor and adjust recruitment strategies (reach) and feedback and goal setting print-out delivery to patients and health-care team members (implementation). Content on print-outs were reinforced by practical webinars providing training about motivational interviewing and collaborative goal setting. The intervention was purposefully brief, low-cost (publicly available), and addressed impact (effectiveness) through standardized assessment and feedback content (23).

Results are summarized elsewhere (24), but in brief, the intervention produced high levels of reach (49% of all eligible patients, including those not contacted), adoption (18 of 30 diverse, low-income clinics approached participated), implementation (all eight risk factors assessed significantly more often in intervention patients; assessment, and print-outs delivered consistently), and effectiveness (intervention superior to randomized paired control clinics on goal setting for 6 of 8 behaviors and changes on 5 of the 8 health behavior and mental health issues). The program was not, however, maintained in any of the settings following conclusion of the study.

To achieve high levels of reach, adoption, and implementation, it was necessary to allow considerable flexibility and customization about how the MOHR was delivered while keeping the content of the intervention standard (23–25). The study was conducted inexpensively and rapidly by the standards of controlled trials.
and demonstrated use of RE-AIM for planning, adaptation, and evaluation. The lack of setting maintenance was due to the inability to integrate the intervention into the existing health records (several different EHR systems were used) and intervention costs while modest (primarily staff time) that exceeded reimbursement provided by Medicare for annual wellness exams.

**Community Setting**

The RE-AIM framework was adopted in the mid-2000s for use by community-based grantees in the aging services and public health networks funded through the Administration for Community Living (26). Use of RE-AIM was part of the grant solicitation, and state grantees were expected to employ RE-AIM in their planning and evaluation of selected evidence-based interventions for managing chronic conditions. RE-AIM was chosen because of its alignment with funder goals: "(1) develop the systems necessary to support the ongoing implementation and sustainability of evidence-based programs for older adults; (2) develop multi-sector community partnerships to enhance program accessibility and extend program capacity; (3) reach the maximum number of at-risk older adults who could benefit from the programs; and (4) deliver evidence-based programs with fidelity" (27). Consultants from the CDC Healthy Aging Research Network (28) provided technical assistance to the grantees (spanning 27 states), who were primarily aging services or public health practitioners, about how RE-AIM elements could be incorporated into their grant processes.

A questionnaire was administered to state grantees to assess the utility of the RE-AIM framework and the integration of RE-AIM elements into different planning, implementation, evaluation, and monitoring processes. Grantees reported RE-AIM was useful for planning, implementation, and evaluation and relevant for various stakeholders (e.g., evaluators, providers, community leaders, and policy makers) (26). For example, RE-AIM influenced grantee decisions about program selection, target populations, and assessment/evaluation tools. Despite the availability of technical assistance, some respondents reported difficulties in use of RE-AIM, especially adopting the framework as a whole. It was not clear if findings reflected grantees’ preferences for adopting single RE-AIM elements over the framework as a whole or if they lacked resources needed to fully assess and track all RE-AIM dimensions.

Over the past decade, RE-AIM utilization has been encouraged in other national-, state-, and local-level community-based initiatives designed to improve the healthy aging. Examples include the CDC’s Initiatives on Assuring Healthy Caregivers (29), Health Foundation of South Florida Healthy Aging Regional Collaborative (30), and the United Way Healthy Aging and Independent Living Initiative (31).

The RE-AIM framework has been valuable for helping community practitioners ask important questions during program planning, implementation, dissemination, and evaluation. However, there is often more use of and adherence to the individual RE-AIM concepts than the model as a whole, which is complicated by the changing lexicon in the field. For example, although the concepts remain consistent, recent federal aging initiatives use terms such as "scalability" and "sustainability" instead of "reach" and "maintenance." Involvement in these aging initiatives reinforces the strong commonality between the study of aging and the RE-AIM framework: both are dynamic processes, evolving over time, and changing with the social context. For continued relevance, frameworks need to be pragmatic, fluid, and adaptable. It is a testimony to RE-AIM that its basic concepts are now mainstreamed and widely integrated into community practice.

**Corporate Setting**

While theoretically as relevant and useful to corporations, the uptake of RE-AIM in corporate settings has been less frequent relative to application in clinical and community settings. Similar to other settings, corporate settings are interested in offering evidence-based programs to their consumers because programs with demonstrated efficacy/effectiveness are most likely to result in positive outcomes, which ultimately satisfies key consumers and stakeholders, and sustains programs (maintenance). Large corporations can have substantial reach because of their infrastructure and support resources (implementation) that enable rapid employment and embedding of the RE-AIM dimensions. This infrastructure allows for systematic program adoption, dissemination, and implementation supported by centralized communication channels and support staff.

The relevance and usefulness of RE-AIM in corporate settings can be demonstrated by closely examining one large US-based corporation, Walgreens. With its 8,175 locations across the US and 87 million rewards account holders, Walgreens has tremendous potential to reach consumers and impact public health. Even a program offered only to Walgreens’ 250,000 employees can have an impact similar to implementing a program to every resident of a moderate-size city.

With an emphasis on trust, care, and accessibility, Walgreens aims to deliver programs that improve its participants’ health and well-being. This is really no different than the goals of many nonprofit, community-based organizations. What is different, however, is that Walgreens’ size and geographic dispersion makes the task of DTI somewhat daunting in terms of logistics and capital needed to initiate a system-wide intervention. Cost and perceived value are the primary reasons that health promotion programs are sustained or discontinued at the community- and corporate-level (Rhodes and Glasgow, unpublished). For example, the incentivized digital health program—Balance Rewards for healthy choices (BRhc)—was implemented in 2014 as a resource-efficient solution to assist Walgreens patients track health behaviors. The value of BRhc has been demonstrated by higher adherence to hypertension and diabetes medications among its users and has shown to promote physical activity among younger adults with chronic conditions (32–34). This program has a vast reach with over one million users, and the digital format of the program moderates the ongoing costs of implementation.

Based on its unique position and infrastructure (like many large corporations), Walgreens has exceeded the capability of many health care and community organizations to deliver an

Rhodes WRD, Glasgow RE. Stakeholder perspectives on costs and resource expenditures: addressing economic issues most relevant to patients, providers and clinics. Unpublished.
intervention with grand-scale reach, adoption, impact, and a maintained presence. However, substantial challenges still exist. Corporations need to value the initial investments and be convinced of adequate return-on-investment for thorough, consistent education and training of delivery staff to achieve reliable results over time (both clinical and financial). If programs are not selected, implemented, and evaluated with the utmost care, the potential patient- and organizational-level loss can be quite damaging. This is a powerful reason to advocate for expanding the application of RE-AIM within corporate settings. Utilizing RE-AIM in corporate settings can produce returns on financial investments while providing benefits to intended populations that are sustained over time.

**DISCUSSION**

This article provided a brief overview of the RE-AIM Framework and its pragmatic use for planning and evaluation while also offering recommendations to facilitate the application of RE-AIM in clinical, community, and corporate settings. Further, this article shared perspectives and lessons learned about employing RE-AIM dimensions in the planning, implementation, and evaluation phases within different settings. Due to nature and restrictions of perspective articles, we focused on limited examples of clinical, community, and corporate work. However, these detailed examples describe initial decision-making, iterative application of RE-AIM processes, and impact on public health outcomes. Similar processes can be applied in other settings for health-related outcomes. Notably, not all evaluations include all RE-AIM dimensions, and there is no right or wrong answer related to which dimensions on which to focus an evaluation. The primary dimensions deserving attention will vary by community, stakeholder and organizational priorities and resources as well as the intervention settings, populations, desired outcomes, and topics. While the processes for reflection and action may differ between clinical, community, and clinical settings based on a unique set of priorities and logistics, the general considerations for applying RE-AIM remain common. To conclude, we discuss lessons learned and recommendations for how RE-AIM can be employed across settings to enhance population health in the future.

A fundamental issue across settings is whether to comprehensively apply the full RE-AIM framework or use a more limited and “strategic” approach to include only certain RE-AIM dimensions. This issue of a full versus pragmatic use of RE-AIM has recently been discussed in detail elsewhere (9, 10, 16, 22, 35), but this topic is especially relevant for applied and unfunded (or underfunded) clinical, community, and corporate non-research settings. For applied settings, the full RE-AIM Framework is best used initially at the outset and planning of a project, and then, select dimensions can be used during and after the program to guide implementation, evaluation, and/or reporting. Initial focus should focus on rough estimates of desired impact for each RE-AIM dimension, followed by decisions about: (A) which dimensions are most important for this project; (B) which dimensions should be measured given limited resources; and (C) which dimensions will be targeted for improvement. This type of pragmatic approach can engage key stakeholders through the use of existing data to determine intervention success (36). A pragmatic approach is intended to allow clinical, community, and corporate settings to consider the entirety of the framework during planning, but then identify actionable RE-AIM information about the most relevant dimensions to determine if a given initiative should be abandoned, refined, sustained, scaled-up, or scaled-out (16).

Given challenges with funding (e.g., more competition to obtain limited resources) in clinical, community, and corporate settings, it is essential to consider strategies to reduce costs and leverage available resources. An interesting concept, frequent need, and important area of study is the “de-implementation” of programs and program elements that appear ineffective, too expensive, or produce unanticipated negative outcomes. Such issues need to be identified in “real time” so an intervention can be quickly modified or discontinued. The urgency of conserving costs and alleviating unnecessary spending (especially at the detriment of community well-being and health equity) highlights the need for ongoing reflection about the RE-AIM dimensions throughout the temporal stages of the intervention. As the RE-AIM framework is used to drive implementation efforts, the same framework can (and should) be used to guide and evaluate de-implementation efforts (37).

A new area of RE-AIM application involves its iterative use to provide ongoing, rapid assessments of progress, then using these results to guide program adaptations (38, 39). For example, early tracking of enrollment (reach) may reveal that key segments of the target population (e.g., low-income patients, those most at risk) are not participating in the intervention. Efforts can then be redirected (and tested) to improve subsequent participation rates. Although RE-AIM was initially used primarily for post hoc program evaluation, it was deemed useful for program planning starting in 2005 (40). Iterative uses of brief, practical measures of targeted RE-AIM dimensions are new and anticipated to grow, which warrants additional research in this area (2).

Our collective experience across clinical, community, and corporate settings indicates the need for greater attention to contextual factors. Often, the most efficient ways to assess contextual factors (the “how and why”) are qualitative or mixed-method approaches (41, 42). Such impressionistic approaches can be helpful to identify conditions under which a program is successful and reasons for such results. The Practical, Robust Implementation, and Sustainability Framework (PRISM) (43) extension of the RE-AIM model may be particularly useful for this purpose because it specifies contextual factor types that may be related to results about different RE-AIM dimensions.

The field of public health has evolved to accommodate changes in societal demographics, the environment, and impacts on the social determinants of health. In fact, such changes have caused new health-related issues and complications that spurred the creation of new fields (e.g., nutrigenomics, computational social science, behavioral economics). As fields advance, so do their needs for sophisticated implementation and evaluation efforts to account for increasing complexity (e.g., big data from multiple sources/levels, nested influence and integrated variables, innovative intervention designs and statistical methodologies, systems issue and unanticipated consequences). We anticipate that the application of RE-AIM will expand to these new fields and offer
within each dimension—as well as potential measures—to bridge interventions across various clinical, community and corporate settings.

AUTHOR CONTRIBUTIONS

All authors contributed to the conceptualization of the manuscript and its content. All authors contributed to the full manuscript as well as reviewed and approved the final version of the manuscript.

FUNDING

The authors would like to acknowledge all members of the National Working Group on RE-AIM Planning and Evaluation Framework (www.re-aim.org). The authors would also like to acknowledge funding support for author contributions: MO and MS contributions supported through ACL SUSTAIN for Better Health and Health Care for Older Adults 90CS0065-01. PE contributions supported by Great Plains IDEA CTR U54 GM115458-01. RE contributions partially supported by IMPlementation to Achieve Clinical Transformation (IMPACT): the Colorado Training Program from the NIH K12 HL137862.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fpubh.2018.00071/full#supplementary-material.

APPENDIX A | Planning and evaluation document, which includes prompts and considerations across all five RE-AIM dimensions by temporal stage within a project (http://www.re-aim.org/resources-and-tools/self-rating-quiz/).

REFERENCES


Conflict of Interest Statement: No financial conflicts of interest to report. All authors are members of the National Working Group on RE-AIM Planning and Evaluation Framework (www.re-aim.org).

Copyright © 2018 Harden, Smith, Orr, Smith-Ray, Estabrooks and Glasgow. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Enhancing the Impact of Implementation Strategies in Healthcare: A Research Agenda

Byron J. Powell, Maria E. Fernandez, Nathaniel J. Williams, Gregory A. Aarons, Rinad S. Beidas, Cara C. Lewis, Sheena M. McHugh and Bryan J. Weiner

The field of implementation science was developed to better understand the factors that facilitate or impede implementation and generate evidence for implementation strategies. In this article, we briefly review progress in implementation science, and suggest five priorities for enhancing the impact of implementation strategies. Specifically, we suggest the need to: (1) enhance methods for designing and tailoring implementation strategies; (2) specify and test mechanisms of change; (3) conduct more effectiveness research on discrete, multi-faceted, and tailored implementation strategies; (4) increase economic evaluations of implementation strategies; and (5) improve the tracking and reporting of implementation strategies. We believe that pursuing these priorities will advance implementation science by helping us to understand when, where, why, and how implementation strategies improve implementation effectiveness and subsequent health outcomes.

Keywords: implementation strategies, implementation science, designing and tailoring, mechanisms, effectiveness research, economic evaluation, reporting guidelines

INTRODUCTION

Nearly 20 years ago, Grol and Grimshaw (1) asserted that evidence-based practice must be complemented by evidence-based implementation. The past two decades have been marked by significant progress, as the field of implementation science has worked to develop a better understanding of implementation barriers and facilitators (i.e., determinants) and generate evidence for implementation strategies (2). In this article, we briefly review progress in implementation science and suggest five priorities for enhancing the impact of implementation strategies. We draw primarily upon the healthcare, behavioral health, and social services literature.
While we hope the proposed priorities are applicable to studies conducted in a wide range of contexts, we welcome discussion regarding potential applications and enhancements for contexts outside of healthcare, such as community and public health settings (3) that often involve different types of stakeholders, interventions, and implementation strategies.

Implementation strategies are methods or techniques used to improve adoption, implementation, sustainment, and scale-up of interventions (4, 5). These strategies vary in complexity, from discrete or single component strategies (6, 7) such as computerized reminders (8) or audit and feedback (9) to multifaceted implementation strategies that combine two or more discrete strategies, some of which have been branded and tested using rigorous designs [e.g., (10, 11)]. Implementation strategies can target a range of stakeholders (12) and multilevel contextual factors across different phases of implementation (13–16). For example, strategies can address patient (17), provider (18), organizational (19), community (20, 21), policy and financing (22), or multilevel (23) factors.

Several taxonomies describe and organize the types of strategies available (6, 7, 24–26). Similarly, taxonomies of behavior change techniques (27) and methods (28) describe components of strategies at a more granular level. Both types of taxonomies promote a common language, inform implementation strategy development and evaluation by facilitating consideration of various “building blocks” or components of multifaceted and multilevel strategies, and improve the quality of reporting in research and practice.

The evidence base for implementation strategies is steadily developing. Initially, single-component, narrowly focused strategies that were effective in earlier studies were selected in subsequent studies despite differences between the clinical problems and contexts in which they were deployed (29). That approach was based on the assumption that strategies would be effective independent of the implementation problems being addressed (29). This “magic bullet” approach has led to limited success (30), prompting recognition that strategies should be selected or developed based upon a thorough understanding of context, including the causes of quality and implementation gaps, an assessment of implementation determinants, and an understanding of the mechanisms and processes needed to address them (29).

Evidence syntheses for discrete, multifaceted, and tailored implementation strategies have been conducted. The Cochrane Collaboration’s Effective Practice and Organization of Care (EPOC) group has been a leader in this regard, with 132 systematic reviews of strategies such as educational meetings (31), audit and feedback (9), printed educational materials (32), and local opinion leaders (33). Grimshaw et al. (34) note that while median absolute effect sizes across implementation strategies are similar (see Table 1), the variation in observed effects within each strategy category suggests that effects may vary based upon whether or not they address determinants (barriers and facilitators). Indeed, determinants at multiple levels and phases may signal the need for multifaceted and tailored strategies that address key determinants (13).

While the use of multifaceted and tailored implementation strategies is intuitive and has considerable face validity (29), the evidence regarding their superiority to single-component strategies has been mixed (37, 39, 40). A review of 25 systematic reviews (39) found “no compelling evidence that multifaceted interventions are more effective than single-component interventions” (p. 20). Grimshaw et al. (34) provide one possible explanation, emphasizing that the general lack of an a priori rationale for the selection of components (i.e., discrete strategies) in multifaceted implementation strategies makes it difficult to determine how these decisions were made. They may have been selected thoughtfully to address prospectively identified determinants through theoretically- or empirically-derived change mechanisms, or they may simply be the manifestation of a “kitchen sink” approach. Wensing et al. (41) offer a complementary perspective, noting that definitions of discrete and multifaceted are problematic. A discrete strategy such as outreach visits may include instruction, motivation, planning of improvement, and technical assistance; thus, it may not be accurate to characterize it as a single-component strategy. Conversely, a multifaceted strategy including educational workshops, educational materials, and webinars may only address provider knowledge and fail to address other important implementation barriers. They propose that multifaceted strategies that truly target multiple relevant implementation determinants could be more effective than single-component strategies (41).

A systematic review of 32 studies testing strategies tailored to address determinants concluded that tailored approaches to implementation were more effective than no strategy or a strategy not tailored to determinants; however, the methods used to identify and prioritize determinants and select implementation strategies were not often well-described and no specific method has been proven superior (37). The lack of systematic methods to guide this process is problematic, as evidenced by a review of 20 studies that found that implementation strategies were often poorly conceived, with mismatches between strategies and determinants (e.g., barriers were identified at the team or organizational level, but strategies were not focused on structures and processes at those levels) (42). A multi-national program of research was undertaken to improve the methods of tailoring implementation strategies (43), but tailored strategies had little impact on primary and secondary outcomes (40). Questions remain about the best methods to develop tailored implementation strategies.

Five priorities need to be addressed to increase the public health impact of implementation strategies: (1) enhance methods for designing and tailoring; (2) specify and test mechanisms of change; (3) conduct more effectiveness research on discrete, multifaceted, and tailored strategies; (4) increase economic evaluations; and (5) improve tracking and reporting. Table 2 provides examples of studies that have pursued each priority with rigor.
we’ve determinants (13–16). Stakeholders may rely upon inertia (i.e., unsystematic manner and fail to address key contextual determinants (13, 14, 15). Thus, strategies are not always well-matched to the contexts in which they are deployed, including the interventions to be implemented, settings, stakeholder preferences, and implementation determinants (37, 42, 54). More rational, systematic approaches to identify and prioritize barriers and link strategies to overcome them are needed (37, 42, 55–57). A number of methods have been suggested. Colquhoun and colleagues (56) found 15 articles with replicable methods for designing strategies to change healthcare professionals’ behavior, and Powell et al. (55) proposed Intervention Mapping (58), concept mapping (59), conjoint analysis (60), and system dynamics modeling (61) as methods to aid the design, selection, and tailoring of strategies. These methods share common steps (identification of barriers, linking barriers to strategy component selection, use of theory, and user engagement), and have potential to make the process of designing and tailoring strategies more rational, systematic, and effective.

TABLE 1 | Evidence for common implementation strategies targeting professional behavior change.

<table>
<thead>
<tr>
<th>Meta-analyses</th>
<th>Number of studies/individuals</th>
<th>Effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Printed educational materials (35)</td>
<td>14 RCTs and 31 ITS</td>
<td>Median absolute improvement of 2.0% (range 0% to 11%)</td>
</tr>
<tr>
<td>Educational meetings (31)</td>
<td>81 RCTs (involving more than 11,000 health professionals)</td>
<td>Median absolute improvement of 6.0% (interquartile range 1.8% to 15.3%)</td>
</tr>
</tbody>
</table>
| Educational outreach (36)                  | 69 RCTs (involving more than 15,000 health professionals) | Median absolute improvements in:  
- Prescribing behaviors (17 comparisons) of 4.8% (interquartile range 3.0–6.5%)  
- Other behaviors (e.g., providing screening tests; 17 comparisons) of 6.0% (interquartile range 3.6–16.0%) |
| Local opinion leaders (33)                 | 18 RCTs (involving more than 296 hospitals and 318 primary care physicians) | Median absolute improvement of care of 12% across studies (interquartile range 6.0–14.5%) |
| Audit and feedback (9)                     | 140 RCTs                      | Median absolute improvement of 4.3% (interquartile range 0.5–16%) |
| Computerized reminders (9)                 | 28 RCTs                       | Median absolute improvement of care 4.2% (interquartile range 0.8–18.8%) |
| Tailored implementation strategies (37)    | 32 RCTs                       | Meta-regression using 15 randomized trials. Pooled odds ratio of 1.56 (95% CI, 1.27–1.93, p < 0.001) |

Table updated from Grimshaw et al. (34), and draws upon Cochrane Reviews from the Effective Practice and Organization of Care (EPOC) group (38).

TABLE 2 | Five priorities for research on implementation strategies.

<table>
<thead>
<tr>
<th>Priority</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Enhance methods for designing and tailoring implementation strategies</td>
</tr>
<tr>
<td></td>
<td>Highfield et al. (44) used Intervention Mapping to systematically design an implementation strategy for an evidence-based mammography intervention.</td>
</tr>
<tr>
<td>2.</td>
<td>Specify and test mechanisms of change</td>
</tr>
<tr>
<td></td>
<td>Williams et al. (45) assessed mechanisms of change for the ARC organizational strategy using multilevel mediation analysis, and found that ARC increases adoption of evidence-based practices by creating proficient organizational cultures that increase clinicians’ intentions to adopt evidence-based practices.</td>
</tr>
<tr>
<td>3.</td>
<td>Conduct more effectiveness research on discrete, multi-faceted, and tailored implementation strategies</td>
</tr>
</tbody>
</table>
|          | Discrete: Gude et al. (46) detail an approach to identifying how a discrete strategy (electronic audit and feedback) works and how it might be optimized.  
Multi-faceted: Kilbourne et al. (47) are using a sequential multiple assignment randomized trial (SMART) design to build an adaptive implementation strategy, demonstrating how trials may be useful in determining the appropriate intensity and sequencing of components in multifaceted strategies.  
Tailored: The Tailored Implementation for Chronic Diseases project provides several examples of trials of tailored implementation strategies and corresponding process evaluations, which are documented in an article collection (48). Wensing details opportunities for improving tailored implementation strategies (49). |
| 4.       | Increase economic evaluations of implementation strategies |
|          | Hoomans and Severens (49) detail how economic analyses apply to implementation science, and provide numerous examples in which cost-effectiveness analyses can facilitate decision-making related to implementation strategies. |
| 5.       | Improve tracking and reporting of implementation strategies |
|          | Tracking: Bunger et al. (50) and Boyd et al. (51) have demonstrated how implementation strategy reporting guidelines (4) can be adapted and applied to prospectively track strategy use over the course of an implementation effort.  
Reporting: Bunger et al. (52) used the Proctor et al. (4) guidelines to retrospectively report key components of a learning collaborative intended to increase the use of Trauma-Focused Cognitive Behavioral Therapy. |

ENHANCE METHODS FOR DESIGNING AND TAILORING IMPLEMENTATION STRATEGIES

Implementation strategies are too often designed in an unsystematic manner and fail to address key contextual determinants (13–16). Stakeholders may rely upon inertia (i.e., “we’ve always done things this way”), one size fits all approaches, or utilize what Martin Eccles has called the ISLAGIATT principle (i.e., “it seemed like a good idea at the time”) (53). Consequently, strategies are not always well-matched to the contexts in which they are deployed, including the interventions to be implemented, settings, stakeholder preferences, and
implementation strategies more rigorous (55, 56). For example, Intervention Mapping is step-by-step approach to developing implementation strategies using a detailed and participatory needs assessment and the identification of implementers, implementation behaviors, determinants, and ultimately, behavior change methods and implementation strategies that influence determinants of implementation behaviors. Some work has been done to compare different methods for assessing determinants (62); however, several questions remain. How can determinants be accurately and efficiently assessed (ideally leveraging implementation frameworks)? Can perceived and actual determinants be differentiated? What are the best methods for prioritizing determinants that need to be proactively addressed? When should determinant assessment take place given that new challenges are likely to emerge during the course of implementation? Who should be involved in this process? Each of those questions has resource implications. Similarly, questions remain about efficiently linking prioritized determinants to effective and pragmatic implementation strategies. How can causal theory be leveraged or developed to guide the selection of implementation strategies? Can pragmatic tools be developed to systematically link strategies to determinants? Approaches to designing and tailoring implementation strategies should be tested to determine whether they improve implementation and clinical outcomes (55, 56). Given that clinical problems, clinical and public health interventions, settings, individuals, and contextual factors are highly heterogeneous, there is much to gain from developing generalizable processes for designing and tailoring strategies.

**SPECIFY AND TEST MECHANISMS OF CHANGE**

Studies of implementation strategies should increasingly focus on establishing the processes and mechanisms by which strategies exert their effects rather than simply establishing whether or not they were effective (29, 63, 64). The National Institutes of Health (64) provides this guidance:

Wherever possible, studies of dissemination or implementation strategies should build knowledge both on the overall effectiveness of the strategies, as well as “how and why” they work. Data on mechanisms of action, moderators, and mediators of dissemination and implementation strategies will greatly aid decision-making on which strategies work for which interventions, in which settings, and for which populations.

Unfortunately, it is not common that mechanisms are even mentioned, much less tested (63, 65, 66). Williams (63) emphasizes the need for trials that test a wider range of multilevel mediators of implementation strategies, stronger theoretical links between strategies and hypothesized mediators, improved design and analysis of multilevel mediation models in randomized trials, and an increasing focus on identifying implementation strategies and behavior change techniques that contribute most to improvement. Developing a more nuanced understanding of mechanisms will require researchers to thoroughly assess the context of implementation and describe causal pathways by which strategies exert their effects, moving beyond a broad identification of determinants and articulating mediators, moderators, preconditions, and proximal and distal outcomes (67). Examples of this type of approach and guidance for their development can be found in Lewis et al. (67), Weiner et al. (23), Bartholomew et al. (58), and Highfield et al. (44). Additionally, drawing more heavily upon theory (66, 68, 69), using research designs that maximize ability to make causal inferences (70, 71), leveraging methods that capture and reflect the complexity of implementation such as systems science (61, 72, 73) and mixed methods (74–76) approaches, and adhering to methods standards for studies of complex interventions (77) will help to sharpen our understanding of how implementation strategies engage hypothesized mechanisms. Work to link implementation strategies and behavior change techniques to hypothesized mechanisms is underway (67, 78), which promises to improve our understanding of how, when, where, and why implementation strategies are effective.

**CONDUCT MORE EFFECTIVENESS RESEARCH ON DISCRETE, MULTI-FACETED, AND TAILORED IMPLEMENTATION STRATEGIES**

There is a need for more and better effectiveness research on discrete, multifaceted, and tailored implementation strategies using a wider range of innovative designs (70, 79–82). First, while a number of discrete implementation strategies have been described (6, 7, 24, 25) and tested (38), there are gaps in our understanding about how to optimize these strategies. There are over 140 randomized trials of audit and feedback, but Ivers et al. (83) conclude that there is much to learn about when it will work best and why, and how to design reliable and effective audit and feedback strategies across different settings and providers. Audit and feedback is an example of how complex implementation strategies can be. The IceBERG group (69) pointed to the fact that even varying five modifiable elements of audit and feedback (content, intensity, method of delivery, duration, and context) produces 288 potential combinations. These variations matter (84), and there is a need for tests of audit and feedback and other discrete implementation strategies that include clearly described components that are theoretically and empirically derived, and well-operationalized. The results of these studies could inform the use of discrete strategies and their inclusion in multifaceted strategies.

Second, there is a need for trials that give insight into the sequencing of multifaceted strategies and what to do if the first strategy fails (39). These strategies could be compared to discrete/single-component implementation strategies or multifaceted strategies of varying complexity and intensity with well-defined components that are theoretically aligned with implementation determinants. These strategies could be tested using MOST, SMART, or other variants of factorial designs that can evaluate the relative impact of
various components of multifaceted strategies and inform their sequencing (70,85).

Finally, tests of strategies that are prospectively tailored to different implementation contexts to address specific implementers, implementation behaviors, or determinants are needed (37). This work could involve comparisons between tailored and non-tailored multifaceted implementation strategies (86), as well as tests of established and innovative methods that could inform the identification, selection, and tailoring of implementation strategies (55, 56).

**INCREASE ECONOMIC EVALUATIONS OF IMPLEMENTATION STRATEGIES**

Few studies include economic evaluations of implementation strategies (87, 88). For example, in a systematic review of 235 implementation studies, only 10% provided information about implementation costs (87). The dearth of economic evaluations severely limits our ability to understand which strategies might be feasible for different contexts, as some decision makers might underestimate the resources required to implement and sustain EBPs, while others might over-estimate them and preemptively limit themselves from implementing EBPs that could benefit their communities (89). Incorporating economic analyses into studies of implementation strategies would provide decision makers more complete information to guide strategy selection, and would encourage researchers to be more judicious and pragmatic in their design and selection of implementation strategies, narrowing attention to strategies and mechanisms hypothesized to be most essential. If methods for designing and tailoring strategies can be improved such that complex multifaceted strategies are proven superior to single-component or less complex multifaceted strategies (39) and tailored strategies are proven superior to more standard multifaceted strategies (37, 40, 43, 55), economic evaluations will be instrumental in demonstrating whether improvements in implementation are worth added costs. Practical tools for integrating economic evaluations within implementation studies have been developed, such as the Costs of Implementing New Strategies (COINS) method (89) which was developed to address the need for standardized methods for analyzing cost data in implementation research that extend beyond the cost of the clinical intervention itself (90). For example, the original COINS study presented a head-to-head trial of two implementation approaches; although one approach was significantly more costly, the implementation outcomes achieved were superior enough to warrant the additional resources (91). Increasing the number and relevance of economic evaluations will require the development of a common framework that promotes comparability across studies (88).

**IMPROVE TRACKING AND REPORTING OF IMPLEMENTATION STRATEGIES**

Developing a robust evidence base for implementation strategies will require that their use be contemporaneously tracked and that they be reported in the literature with sufficient detail (92). It is often difficult to ascertain which implementation strategies were used and how they might be replicated. Part of the challenge is the iterative nature of implementation. Even if strategies are meticulously described in a study protocol or trial registry, it is often unrealistic to expect that they will not need to be altered as determinants emerge across implementation phases (13, 93, 94). These changes are likely to occur within and between implementing sites in research studies and applied efforts (50, 51), and without rigorous methods for tracking implementation strategy use, efforts to understand what strategies were used and whether or not they were effective are stymied. Even when strategies are reported in study protocols or empirical articles, there are numerous problems with their description, including inconsistent labeling; lack of operational definitions; poor description and absence of manuals to guide their use; and lack of a clear theoretical, empirical, or pragmatic justification for how the strategies were developed and applied (4). Poor reporting clouds the interpretation of results, precludes replication in research and practice, and limits our ability to synthesize findings across studies (4, 92). Findings from systematic reviews illustrate this problem. For example, Nadeem et al. (95) review of learning collaboratives concluded that, “reporting on specific components of the collaborative was imprecise across articles, rendering it impossible to identify active quality improvement collaborative ingredients linked to improved care.”

A number of reporting guidelines could be leveraged to improve descriptions of strategies (4, 96–100). Proctor et al. (4) recommend that researchers name and define strategies in ways that are consistent with the published literature, and carefully operationalize the strategy by specifying: (1) actor(s), (2) action(s), (3) action target(s), (4) temporality, (5) dose, (6) implementation outcomes affected, and (7) theoretical, empirical, or pragmatic justification. Specifying strategies in this way has the potential to increase our understanding of not only which strategies are most effective, but more importantly, the processes and mechanisms by which they exert their effects (29, 67). Additional options that provide structured reporting recommendations include the Workgroup for Intervention Development and Evaluation Research (WIDER) recommendations (99, 100), the Simplified Framework (96) and its extension [AIMD; (97)], and the Template for Intervention Description and Replication (TIDieR) checklist (98). Though not specific to the reporting of implementation strategies, the Standards for Reporting Implementation Studies (101) and Neta et al. (102) reporting framework emphasizes how critical it is to report on the multilevel context of implementation. The use of any of the existing guidelines would enhance the clarity of strategy description. We believe that developing approaches to tracking implementation strategies (50, 51), and assessing the extent to which they are pragmatic (e.g., acceptable, compatible, easy, and useful) for both research and applied efforts is a high priority. Further, efficient ways of linking empirical studies with study protocols to gauge the degree to which strategies have been adapted or tailored over the course of an implementation
Implementation science has advanced considerably, yielding a more robust understanding of implementation strategies. Several resources can inform the use of implementation strategies, including established taxonomies of implementation strategies (6, 7, 24, 25) and behavior change techniques (27, 28), repositories of systematic reviews (38, 103, 104), methods for selecting and tailoring implementation strategies (40, 55, 56), and reporting guidelines that promote replicability (4, 98–100). Nevertheless, questions remain and further effectiveness research and methodological development are needed to ensure that evidence is effectively translated into public health impact. Advancing these priorities will lead to a better understanding of when, where, why, and how implementation strategies exert their effects (29, 63).

CONCLUSION


REFERENCES


Powell et al. Enhancing the Impact of Implementation Strategies


methodology-standards?utm_source=Funding+awards%2C+GAO+Board+deadline&utm_campaign=Funding+awards%2C+GAO+Board+deadline&utm_medium=email#Complex


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Powell, Fernandez, Williams, Aarons, Beidas, Lewis, McHugh and Weiner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
The Price per Prospective Consumer of Providing Therapist Training and Consultation in Seven Evidence-Based Treatments within a Large Public Behavioral Health System: An Example Cost-Analysis Metric

Kelsie H. Okamura1,2*, Courtney L. Benjamin Wolk1, Christina D. Kang-Yi1, Rebecca Stewart1, Ronnie M. Rubin3, Shawna Weaver3, Arthur C. Evans4, Zuleyha Cidav1, Rinad S. Beidas1 and David S. Mandell1

1 Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, United States, 2 State of Hawaii Child and Adolescent Mental Health Division, Honolulu, HI, United States, 3 City of Philadelphia Department of Behavioral Health and Intellectual Disability Services, Philadelphia, PA, United States, 4 American Psychological Association, Washington, DC, United States

Objective: Public-sector behavioral health systems seeking to implement evidence-based treatments (EBTs) may face challenges selecting EBTs given their limited resources. This study describes and illustrates one method to calculate cost related to training and consultation to assist system-level decisions about which EBTs to select.

Methods: Training, consultation, and indirect labor costs were calculated for seven commonly implemented EBTs. Using extant literature, we then estimated the diagnoses and populations for which each EBT was indicated. Diagnostic and demographic information from Medicaid claims data were obtained from a large behavioral health payer organization and used to estimate the number of covered people with whom the EBT could be used and to calculate implementation-associated costs per consumer.

Results: Findings suggest substantial cost to therapists and service systems related to EBT training and consultation. Training and consultation costs varied by EBT, from Dialectical Behavior Therapy at $238.07 to Cognitive Behavioral Therapy at $0.18 per potential consumer served. Total cost did not correspond with the number of prospective consumers served by an EBT.

Conclusion: A cost-metric that accounts for the prospective recipients of a given EBT within a given population may provide insight into how systems should prioritize training efforts. Future policy should consider the financial burden of EBT implementation in relation to the context of the population being served and begin a dialog in creating incentives for EBT use.

Keywords: evidence-based treatment, therapist, training and consultation, cost-analysis, population health
INTRODUCTION

In recent years, many efforts to improve mental health have focused on increasing the use of evidence-based treatments (EBTs) within public-sector service systems. Therapist training is a necessary—but not sufficient—implementation strategy to increase EBT use (1). For public-sector service systems, large-scale training of therapists is often the first or only EBT implementation strategy. A combination of experiential and active learning (e.g., didactic and case consultation) tends to produce the most favorable therapist behavior change over time (2, 3). As a result, many EBT developers and certifying organizations now require that therapists receive both didactic foundational training and ongoing consultation to be “certified” in an EBT (e.g., PCIT International).¹ Training and consultation require an investment on the part of therapists, their agencies, and, especially in publicly funded systems, the city or state agency that oversees payment for care. For example, therapists and organizations may incur initial direct costs like attending week-long trainings to first learn about the EBT and subsequently participate in weekly consultation calls for 6–12 months to ensure treatment fidelity. Therapists’ time required to participate often results in substantial cost to the agencies which they work (4–8). For example, Lang and Connell (6) estimated that an agency participating in a Trauma Focused-Cognitive Behavioral Therapy learning collaborative, which included agency-wide training and ongoing consultation, spent $89,575 in direct (e.g., training) and indirect (e.g., preparation hours) costs.

While public-sector service systems have typically used other strategies to select EBT, such as stakeholder feedback in combination with federal- and/or state-policy (9, 10), the breadth of the population served, and the associated costs should be important drivers of choice. Utilizing existing service system data is important for strategic decision-making and implementation tailored to the population (11). Information about the population served is needed to make decisions about where to invest their limited resources by understanding the extent to which an EBT provides diagnostic and demographic “coverage” within a service system (9). Costs associated with EBTs are often noted as significant barriers for implementation (12–14) and thus far the cost-analysis metrics that have been used to study implementation have not considered the population coverage relative to the implementation cost (4–8). A metric that considers the potential consumer served allows for population-based and data-informed decisions when selecting the right EBT. This metric can also inform cost- evaluative decisions on how applicable an EBT will be for each relevant consumer within the service system.

In this study, we introduce a strategy for calculating a cost per prospective consumer metric to determine the extent to which an EBT covers a given service system. To generate this metric, population data derived from that existing service system are needed; and within behavioral health, insurance claims (15), or practice-monitoring data tied to billing (9, 16) have predominantly been used. These large person-period datasets typically contain information regarding consumer age, gender, diagnoses, service utilization, and medication prescribed. This study was conducted to demonstrate the impact of therapist training and consultation costs in a large public behavioral health system and to describe a complimentary metric for system decision-making when selecting EBT for their population. First, training and consultation requirements for certification among seven EBTs were documented. Next, training, consultation, and indirect labor costs for each EBT were calculated. Finally, the total cost of training, consultation, and indirect labor for each EBT was divided across the number of potential consumers based on diagnostic and demographic information.

MATERIALS AND METHODS

Evidence-Based Treatments

Identification of EBT

We identified EBTs for this study using registries created by the American Psychological Association (17), which rely on Chambless and Hollon (18) definitions of EBT. The APA’s Division 12 (Society for Clinical Psychology)² and Division 53 (Society for Child and Adolescent Clinical Psychology)³ websites were consulted to determine EBTs that fit the criteria of (a) having an in-person training, (b) ongoing consultation period, and (c) a certifying body through which therapists can become “certified” in the particular EBT. Seven EBTs were identified through these websites: (a) Cognitive Behavioral Therapy/Cognitive Therapy (19), (b) Cognitive Processing Therapy (20), (c) Dialectical Behavior Therapy (21), (d) Parent–Child Interaction Therapy (22, 23), (e) Prolonged Exposure (24, 25), (f) Modular Approach to Therapy for Children with Anxiety, Depression, Trauma, and Conduct Problems (26), and (g) Trauma Focused-Cognitive Behavioral Therapy (27).

Training and Consultation Cost

The cost of training and consultation was determined using information from the certifying body for each EBT (see Table 1 for certifying bodies for each EBT). First, the certifying body’s website was referenced for certification requirements, upcoming trainings, and cost associated with training, consultation, and certification. When prices were not listed, we contacted the certifying body to solicit current prices and requirements for training and consultation to obtain certification. Revenue loss was defined as the total amount of therapist hours spent on training and consultation, as opposed to providing therapy (i.e., billable hours). Hourly wage for therapists, as determined by the US Bureau of Labor Statistics for Philadelphia,⁴ was established as $38.37 per hour.

Diagnostic and Age Applicability

To determine the population to which an EBT was applicable, diagnostic and age profiles were created for each EBT. We referenced APA’s Divisions 12 and 53 websites, the credentialing body’s website, and PracticeWise Evidence-based Services Database (28) to identify the studies used to establish each EBT’s

¹http://www.pcit.org/.
³http://effectivechildtherapy.org/.
efficacy. For example, Division 12’s website lists DBT as having Strong Research Support for Borderline Personality Disorder, with six efficacy trials used to determine that status. The Division 12 website also lists Strong Research Support for CBT/CT for Attention-Deficit/Hyperactivity Disorder, Insomnia, Eating Disorder, Bipolar Disorders, Bulimia Nervosa, Depressive Disorders, Generalized Anxiety Disorder, Obsessive Compulsive Disorder, Social Phobia, Panic Disorder, and Schizophrenia; CPT for Posttraumatic Stress Disorder; and PE for Posttraumatic Stress Disorder. Efficacy trials for PCIT, MATCH, and TF-CBT were identified through comprehensive literature reviews cited by Division 53 (29–31) and the credentialing body’s website (PCIT International, PracticeWise, and TF-CBT National Therapist Certification Program, respectively).

Efficacy trials were coded by two independent raters (Kelsie H. Okamura and Courtney L. Benjamin Wolk) for diagnosis and age range used within each trial. Coders met to regularly resolve discrepancies, using clinical judgment and the conservative criteria of only including diagnoses that the EBT was intended to treat. Specifically for youth CBT, the PracticeWise Evidence-based Services Database (32, 33), a searchable database synthesizing more than 800 treatment studies for youth with psychiatric disorders, was referenced to determine a CBT youth diagnostic and age profile. The database was searched for CBT trials to identify diagnoses and age ranges that met well-established criteria proposed by Chambless and Hollon (18).

**Population-Based Data Source and Study Sample**

Philadelphia County behavioral health Medicaid claims ($N = 903,980$) were used to identify a subset of consumers ($N = 60,391$) who received outpatient behavioral health services during November 2015 through October 2016. This 1-year time period was chosen because of the shift from ICD-9 and DSM-IV-TR diagnoses to ICD-10 and DSM-5 diagnoses. De-identified claims included age at the first claim, sex, race, psychiatric diagnosis, and behavioral health service use. Behavioral health services were categorized based on level of care codes and only claims reflective of outpatient therapy services were retained (i.e., assessment and medication management codes were excluded). The final sample included the consumers with two or more outpatient claims aggregated by ICD-10 diagnosis. Consumers may have been counted more than once across but not within ICD-10 diagnoses. This allowed for more consumer coverage and the ability to account for multiple psychiatric diagnoses. The University of Pennsylvania and the City of Philadelphia Department of Public Health Institutional Review Boards determined that this study was exempt from review due to the masking of identifiable information.

The final study sample included 897,064 claims representing 53,475 unique consumers. There were 6,916 duplicate consumers removed from analyses due to multiple claims being submitted for the same consumer for more than one diagnosis. In instances of multiple claims, the first claim per consumer was retained. Consumers were 53.4% female ($n = 34,507$) and averaged 29.91 (SD = 17.99) years of age. Race included African-American (42.7%, $n = 27,573$), Hispanic (37.8%, $n = 24,339$), White (15.6%, $n = 10,061$), and Other (3.9%, $n = 2,531$).

**Cost-Analysis Metric**

The cost of therapist training, consultation, certification, and revenue loss were summed to calculate a total training and consultation cost for each EBT. This total training and consultation therapist cost was then divided by the number of consumers within Philadelphia County Medicaid claims who matched the

---

EBT diagnostic and age profile. This formula resulted in an EBT training and consultation cost per potential consumer:

\[
\text{EBT Training and consultation cost per potential consumer} = \frac{\text{TRAINING + CONSULTATION} + \text{CERTIFICATION} + \text{REVENUE LOSS}}{\text{NUMBER OF PROSPECTIVE CONSUMERS SERVED}}.
\]

RESULTS

Training and Consultation Requirements
Certifying bodies, training hours, consultation hours, training length, and specific criteria related to consultation are detailed in Table 1. Across EBTs, 2–5 days of in-person training were required for certification. TF-CBT and CPT both required online training in addition to the in-person training. Trainings were provided by certified trainers in each respective EBT, identified by the certifying body. Regarding consultation, DBT and PCIT required the most ongoing consultation (i.e., bimonthly contact for approximately a year), whereas CBT/CT required fewer hours (i.e., 1 year of clinical experience with 10 h of consultation). Live feedback in the form of tape review or telehealth observation was included in the consultation descriptions for PCIT and PE. Consultation hours typically spanned 6–12 months. MATCH and TF-CBT gave the option of meeting twice per month for 6 months or once per month for 12 months. PCIT and PE consultation were based on completion of two cases rather than a set time frame. CPT was similar in that it required 20 h of group or 12 h of individual consultation. Consultation was provided by a certified supervisor identified by the certifying body.

Training and Consultation Cost
Training, consultation, certification, and revenue loss costs were summed to form a total cost in Table 2. EBT are rank ordered by their total cost, with DBT being the most expensive to TF-CBT being the least expensive. Training costs ranged from $585 for CPT to $4,900 for PCIT per therapist. However, consultation costs are included in the PCIT training cost. In addition to PCIT, MATCH and TF-CBT included the cost of consultation into their training cost. Stand-alone consultation prices ranged from $2,000 to $12,500, with consultation costs as either a set rate (i.e., $2,000 for CBT/CT consultation), per session rate (i.e., $185 for PE), or an hourly rate (i.e., $250 per hour for DBT, $200 per hour for CPT).

Cost per Prospective Consumer
Prospective consumer costs were calculated by summing the total cost of training, consultation, certification, and revenue loss, and dividing that among the number of unique consumers fitting each EBT diagnostic and age profile. Table 3 details the total cost, age range in years, diagnoses, number of unique consumers fitting the diagnostic and age profile, and a cost per consumer (total cost/consumers) and is ordered by the per prospective consumer cost (most to least expensive). Cognitive Behavioral Therapy/Cognitive Therapy was the least expensive per consumer ($0.18) and covered the most prospective consumers (n = 39,586). In contrast, DBT was the most expensive per consumer ($238.07) and covered the fewest prospective consumers (n = 81).

DISCUSSION

The goal of this study was to develop a cost-analysis metric around the specific implementation strategy of EBT training and consultation while considering the population being served. This is particularly important given the financial pressures that large behavioral health services systems face to effectively implement EBT and manage tax-payer dollars and costs to the system, agencies, therapists, and consumers. Our study used seven common EBTs and compared training and consultation hours and prices and calculated per prospective consumer costs in a large behavioral health system. Training and consultation requirements and costs varied widely across EBT. Training and consultation costs ranged from $600 to $14,985 per therapist, and when considering certification fees and revenue loss from time spent in training rather than serving consumers, total costs ranged from $2,231.32 to $19,283.30. This represents a substantial investment to therapists, organizations, and systems. For some EBTs, consultation emerged as the most time-consuming and costly aspect, which is often emphasized as an important implementation strategy (2). Total cost did not correspond with the number of prospective consumers served by an EBT in our current behavioral health system sample. That is, the most expensive EBTs were not those that the most prospective consumers would benefit. This cost-analysis metric utilizing prospective consumer behavioral health outpatient claims appears to be a useful tool for large system decision-making in choosing EBT.

The costliest EBT to train (i.e., DBT) covered the fewest consumers in the system, likely because few consumers had

<table>
<thead>
<tr>
<th>EBT</th>
<th>Tuition</th>
<th>Consultation</th>
<th>Training and consultation</th>
<th>Certification</th>
<th>Revenue loss</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBT</td>
<td>$2,485.00</td>
<td>$12,500.00</td>
<td>$14,985.00</td>
<td>$845.00</td>
<td>$3,453.30</td>
<td>$19,283.30</td>
</tr>
<tr>
<td>PCIT</td>
<td>$4,900.00</td>
<td>Included</td>
<td>$4,900.00</td>
<td>$225.00</td>
<td>$3,453.30</td>
<td>$8,578.30</td>
</tr>
<tr>
<td>PE</td>
<td>$1,500.00</td>
<td>$3,885.00</td>
<td>$5,385.00</td>
<td>Included</td>
<td>$2,033.61</td>
<td>$7,416.81</td>
</tr>
<tr>
<td>CBT/CT</td>
<td>$2,700.00</td>
<td>$2,000.00</td>
<td>$4,700.00</td>
<td>$450.00</td>
<td>$1,918.50</td>
<td>$7,068.50</td>
</tr>
<tr>
<td>CPT</td>
<td>$585.00</td>
<td>$2,585.00</td>
<td>$2,585.00</td>
<td>$250.00</td>
<td>$1,381.32</td>
<td>$4,216.32</td>
</tr>
<tr>
<td>MATCH</td>
<td>$1,900.00</td>
<td>$1,900.00</td>
<td>$3,800.00</td>
<td>$158.00</td>
<td>$1,995.24</td>
<td>$4,053.24</td>
</tr>
<tr>
<td>TF-CBT</td>
<td>$600.00</td>
<td>Included</td>
<td>$600.00</td>
<td>$250.00</td>
<td>$1,381.32</td>
<td>$2,231.32</td>
</tr>
</tbody>
</table>

EBT, evidence-based treatment; DBT, Dialectical Behavior Therapy; PCIT, Parent–Child Interaction Therapy; PE, Prolonged Exposure; CBT/CT, Cognitive Behavioral Therapy/Cognitive Therapy; CPT, Cognitive Processing Therapy; MATCH, Modular Approach to Therapy for Children with Anxiety, Depression, Traumas, or Conduct Problems; TF-CBT, Trauma Focused-Cognitive Behavior Therapy.
a borderline personality disorder diagnosis. It is important to reiterate here that we used conservative diagnostic criteria for classifying which disorders a treatment was evidence-based for, and as such, may have excluded groups of consumers that may benefit from DBT (e.g., youth with suicidal ideation). We discuss this more in our limitations section as well as the cost-savings of having such a specialized EBT within a behavioral health system. Some of the less expensive EBTs provided greater consumer coverage. Systems considering which EBTs to invest in may wish to consider a tiered approach. That is, begin with (a) a generalist EBT (i.e., CBT/CT and MATCH) and then consider adding on (b) trauma focused (i.e., TF-CBT, PE, or CPT), and (c) other specialty EBT (i.e., PCIT and DBT) depending on the prospective consumers served. The proposed cost-analysis metric may be particularly useful for systems seeking to understand the financial impact of specialty EBT (34). While most costly in our study, if a specialty EBT like DBT aligns well with system priorities, such as reducing inpatient hospitalization rates, residential treatment utilization, or other out of home placement, it may make the additional investment worthwhile. Furthermore, it may be beneficial for systems to create a ratio of therapists trained to prospective consumers served to inform future training efforts. This tiered approach also has implications for research which is beginning to suggest that attitudes (35) and knowledge (36) vary by practices and EBT, suggesting that our field’s conceptualization of EBT as all-encompassing may be misguided. Moreover, treatment developers may wish to consider building modularity and tiered decision-making into interventions to increase applicability to a broader range of consumers. A tiered approach to choosing and conceptualizing EBT may facilitate decisions about which EBTs to compare and study within effectiveness and implementation studies (e.g., comparing two generalist type EBTs rather than a specialty EBT and generalist EBT).

In-person didactic training and ongoing consultation were required across all seven EBTs for certification. The typical time period for in-person training was 1 week (40 h); however, CPT and TF-CBT required only 2 days (24 h) in-person training with completion of an additional online course as a pre-requisite for certification. Reviews of empirical studies on training have concluded that didactic training alone does not produce change in therapist behavior and should be combined with ongoing feedback and consultation (2, 3). However, it is unclear from the literature to which extent didactic trainings need to be delivered in-person and the requisite amount of training hours to attain competency. Our findings suggest an emerging standard of 40 h for didactic training. From a system’s perspective, taking cohorts of service-delivering therapists offline for a week may be perceived as both costly and detrimental to consumers receiving services. However, if multiple systems begin to adopt this convention of training and consultation as requirements for employment and credentialing as well as enhance outpatient rates to absorb some of those costs, they may be more acceptable and feasible to provider agencies.
Ongoing consultation requirements also place considerable demand on the therapist and system. In this study, consultation requirements were observed to vary even more than didactic training requirements. For example, CPT, MATCH, and TF-CBT required 12 h of supervision across varying time frames (e.g., 6–12 months, see Table 1), whereas DBT and PCIT required a year of ongoing consultation with bimonthly attendance. Research has suggested that the purpose of ongoing consultation is to give the trainee the opportunity to apply the skills learned in didactics with sufficient supervision and support (37, 38). Typically, consultation entails ongoing case-review, which may or may not take the form of reviewing session recordings or live feedback. Indeed, only PCIT and PE included live or taped feedback as a part of their consultation model. Consistent with didactic training, the frequency, and depth of consultation needed to fully achieve competency has not been established and this may impact cost. For example, consultation with review of session recordings is more time-consuming than case-based discussions. Furthermore, research on training and sustainability has noted that even when therapists are comprehensively trained and supervised in EBT they do not use EBT frequently in their practice (2). Determining the optimal duration and format for didactic training and consultation should be an implementation science priority. For public-sector service systems, there are likely many considerations when deciding which EBT(s) to invest in including time, cost, policy, and population-based characteristics. For example, should a service system first choose an EBT that requires less training and consultation (e.g., MATCH) over one that requires a longer training and consultation time frame (e.g., DBT) to increase EBT capacity quickly? The answer to this question is beyond the scope of this study. However, initial findings suggest that the variation between EBTs is substantial enough to warrant further attention.

The results of this study should be considered within the context of several limitations. First, our study used administrative Medicaid claims data, which may not be reflective of the entire service-seeking population (e.g., private insurance covered consumers or population prevalence within the community). Furthermore, several studies have suggested that Medicaid claims data may not be diagnostically accurate (39–41). However, studies have demonstrated that the agreement of Medicaid claims diagnoses to clinical data is around 85% (39, 40) suggesting that the inaccuracy of claims may be related to an under-identification of disorders rather than inaccuracy of diagnosis. Also related to diagnosis, some of the efficacy trials that we coded to create age and diagnostic profiles included multiple psychiatric diagnoses, which may suggest that the corresponding EBT would be appropriate for both the intended and comorbid conditions. In these instances, we took a conservative approach and only considered the diagnoses for which an EBT primarily targeted. For example, a trial of DBT for individuals with borderline personality disorder and cooccurring substance use was coded as effective for adults with a diagnosis of borderline personality disorder but not for individuals with a primary substance use disorder diagnosis. Future studies may wish to examine broader diagnostic categories (e.g., depressive disorders versus major depressive disorder) or behavioral codes (e.g., suicidality), which may represent a more inclusive approach. In addition, replication using national epidemiological data with standardized diagnostic assessments [e.g., Ref. (42, 43)] would circumvent concerns about diagnostic accuracy and provide additional insight into the proportion nationally that might benefit from specific EBTs. It is also important to note that this cost-analysis metric, while not statistically or methodologically difficult to apply, does require some expertise in using claims data. Therefore, public-sector service systems will need administrators, analysts, or external research/evaluation partners to apply the cost-analysis metric to claims datasets.

We examined the costs associated with specific implementation strategies (i.e., training and consultation) without considering the effectiveness of the intervention itself (i.e., cost-effectiveness analysis, especially in the case of DBT). Raghavan (44) has noted that estimating implementation costs is different from cost-effectiveness as it is influenced by the entity (e.g., system, agency, and therapist) to which the cost is associated as well as the strategy, EBT, and setting (45). Our goal was to understand the direct and indirect costs at the population level that may be associated with the implementation strategy of EBT training and consultation in a large public behavioral health system. One important caveat was that training and consultation costs were calculated at the individual therapist level, which may not parallel costs for system-wide trainings in the community (46). Often, partnerships and contracts are executed to train and provide consultation for large cohorts of therapists within the system versus using a cost per therapist model (47). In addition, indirect costs were calculated based on therapist wage loss during training and consultation (and not revenue loss to the provider agency), without accounting for other contributing activities to sustaining the EBT including supervision, non-billable preparation hours, and travel time. Again, our focus was on the implementation strategy of training and consultation and is consistent with other studies that have evaluated a discrete amount of time as a part of the indirect implementation cost (44). Furthermore, Beidas et al. (12) have demonstrated that high turnover often affects the fiscal landscape of EBP implementation and our study did not account for loss on investment or the extent to which a therapist needed to stay within the system for a good return on investment. System policy makers, administrators, and researchers will need to collaboratively set standards for training requirements and cost and conduct cost-effectiveness studies that are linked to consumer outcomes.

Despite these limitations, this study proposes a methodology for considering which EBT to choose within a large behavioral health system. We propose a tiered approach to selecting EBT, allowing our cost-analysis metric, stakeholder feedback, and system priorities to influence the selection. Our cost calculations may also serve as a basis for policy around incentivizing the use of EBT (1), especially in the early stages of implementation when the system and agency can expect a loss in revenue due to therapist productivity and agency revenue. For example, Timmer and Urquiza (48) described a demonstration project in Los Angeles County Department of Mental Health that reimbursed agencies for lost productivity hours during an initial training initiative. While some systems have mandated the use of EBT (49, 50), few
systems have begun to incentivize the use of EBT (i.e., Chester County, PA, USA; City of Philadelphia Department of Behavioral Health and Intellectual disAbility Services). Understanding the effectiveness of mandates and incentives in therapist utilization and consumer receipt of EBT as well as improved clinical outcomes will be the next era of implementation research, and developing pragmatic cost-analysis metrics will enable large systems to make decisions about which EBT to adopt for whom. Moreover, developing methods and testing them within and across large systems of care will enhance implementation science and generalizability of findings in health services research.

**AUTHOR CONTRIBUTIONS**

KO was responsible for all aspects of this manuscript, from conceptualization to writing. CW and DM provided consultation and editing. CK-Y performed data analysis of Medicaid claims data. ZC provided consultation in health economics. RS, RB, RR, SW, and AE provided additional feedback and editing of the manuscript.

**FUNDING**

KO is a 2017 recipient of the Child Intervention, Prevention, and Services (CHIPS) Fellowship, funded through an award from the National Institute of Mental Health (5R25MH06836713) and a Robert Wood Johnson Foundation New Connections Scholar. CW is an investigator with the Implementation Research Institute (IRI), at the George Warren Brown School of Social Work, Washington University in St. Louis; funded through an award from the National Institute of Mental Health (5R25MH08091607) and the Department of Veterans Affairs, Health Services Research & Development Service, Quality Enhancement Research Initiative (QUERI). RS (F32MH103960) and RB (K23MH099179) receive research support through the National Institute of Mental Health. RB, ZC, DM, KO, and CW are fellows of the Leonard Davis Institute for Health Economics, University of Pennsylvania.

**REFERENCES**

train-the-trainer model for youth cognitive behavior therapy. Developing a systematic evaluation approach for training programs within a settings.


Conflict of Interest Statement: No authors declare any personal, professional, or financial relationships that could potentially be construed as a conflict of interest.

The reviewer AB and handling Editor declared their shared affiliation.

Copyright © 2018 Okamura, Benjamin Wolk, Kang-Yi Stewart, Rubin, Weaver, Evans, Cidav, Beidas and Mandell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Unpacking Partnership, Engagement, and Collaboration Research to Inform Implementation Strategies Development: Theoretical Frameworks and Emerging Methodologies

Keng-Yen Huang 1*, Simona C. Kwon 1, Sabrina Cheng 1, Dimitra Kamboukos 1, Donna Shelley 1, Laurie M. Brotman 1, Sue A. Kaplan 1, Ogedegbe Olugbenga 1 and Kimberly Hoagwood 2

1 Department of Population Health, New York University School of Medicine, New York, NY, United States, 2 Child and Adolescent Psychiatry, New York University School of Medicine, New York, NY, United States

Background: Partnership, engagement, and collaboration (PEC) are critical factors in dissemination and implementation (D&I) research. Despite a growing recognition that incorporating PEC strategies in D&I research is likely to increase the relevance, feasibility, impacts, and of evidence-based interventions or practices (EBIs, EBPs), conceptual frameworks and methodologies to guide the development and testing of PEC strategies in D&I research are lacking. To address this methodological gap, a review was conducted to summarize what we know, what we think we know, and what we need to know about PEC to inform D&I research.

Methods: A cross-field scoping review, drawing upon a broad range of PEC related literature in health, was conducted. Publications reviewed focused on factors influencing PEC, and processes, mechanisms and strategies for promoting effective PEC. The review was conducted separately for three forms of partnerships that are commonly used in D&I research: (1) consumer-provider or patient-implementer partnership; (2) delivery system or implementation team partnership; and (3) sustainment/support or interagency/community partnership. A total of 39 studies, of which 21 were review articles, were selected for an in-depth review.

Results: Across three forms of partnerships, four domains (cognitive, interpersonal/affective, behavioral, and contextual domains) were consistently identified as factors and strategies for promoting PEC. Depending on the stage (preparation or execution) and purpose of the partnership (regulating performance or managing maintenance), certain PEC strategies are more or less relevant. Recent developments of PEC frameworks, such as Partnership Stage of Change and multiple dynamic processes, provide more comprehensive conceptual explanations for PEC mechanisms, which can better guide PEC strategies selection and integration in D&I research.

**Conclusions:** This review contributes to D&I knowledge by identifying critical domain factors, processes, or mechanisms, and key strategies for PEC, and offers a multi-level PEC framework for future research to build the evidence base. However, more research is needed to test PEC mechanisms.

**Keywords:** engagement, collaboration, partnership, patient engagement, patient-centered, community engagement, team science, implementation strategies

**BACKGROUND**

**Introduction**

Dissemination and Implementation (D&I) research, which involves the use of diverse strategies to facilitate adoption, integration, and sustainability of evidence-based interventions and practices (EBIs/EBPs) in diverse settings, is a rapidly growing field in health research (1). The successful development, implementation, dissemination, and sustainability of EBIs/EBPs requires communication, collaboration, and consensus among all involved, including consumers (end users), implementers, and related partners who contribute to sustainability of EBIs/EBPs.

To accomplish this, Partnership, partner Engagement, and Collaboration (PEC) have been identified as critical strategies in D&I research (2, 3). In a recent compilation of recommended strategies by D&I experts, more than one-third were PEC-related strategies (e.g., coalition building, creating a learning collaborative, developing academic partnerships, involving patient/consumers and family members, organizing clinician implementation meetings, and promoting network weaving) (3).

Despite the importance of PEC strategies and their potential contribution for improving service implementation and health outcomes at the individual, community, and population levels, the conceptualization and methodologies for studying PEC are not well defined and not well integrated into D&I research. Although PEC research has been applied in multiple fields over the years, including military, business, sports, academia, and health, lessons learned, and findings from these fields have not been systematically applied to inform PEC strategies for enhancing EBIs/EBPs in implementation research.

**Multilevel Partnership, Engagement, and Collaboration in D&I Research**

In D&I research, PEC can be applied across different programs and interventions (4) and with diverse partners form multiple levels (5). Partners involved in D&I usually include consumers, a team of providers/implementers (i.e., those who provide EBIs/EBPs), and a team of multi-disciplinary partners (i.e., those who set up structures and policies, and provide support for implementation and sustainment of EBIs/EBPs). The purpose of developing strong PEC in D&I research is to build support across the individual, team, and organizational levels to work toward common goals for EBIs/EBPs, use or share skills and resources to implement EBIs/EBPs, and seek input and support of experts from different disciplines. Therefore, D&I research requires consideration of PEC strategies for multiple forms or multiple levels of partnerships.

At the consumer-provider level (consumers also defined as patients or targets of EBIs/EBPs, and providers also defined as implementers who provide EBIs/EBPs), PEC between consumers and providers is critical because substantial research has documented that an effective patient-provider relationship can optimize the patient’s use of intervention strategies and engagement in treatment (6, 7). Greater patient-centered care or patient-provider partnerships are associated with better patient outcomes, including increased health knowledge, management skills, competency, self-efficacy, and sense of control and wellbeing over personal health, and well-being (8). Additionally, better patient-provider partnership also benefits providers because of increased patient satisfaction with their care (8).

Therefore, application of PEC strategies to promote patient-provider partnerships has implications to improve EBIs/EBPs acceptability and patient-centered care outcomes, and to enhance patients’ use of EBI health promotion and management strategies.

At the EBI/EBP delivery system level (or implementation team level), quality of interaction, relationship and behavioral processes of implementation team members can influence teams’ performance and effectiveness in EBIs/EBPs implementation (9, 10). Partnership and implementation barriers that are commonly identified at this level include: lack of effective communication and coordination among team members, lack of sufficient buy-in from team members, high turnover, failure of partnership leaders to engage team members, lack of sufficient funds to support partnerships, and team member burnout (11). Therefore, PEC strategies that engage teams’ long term collaborative efforts, and empower and motivate members to proactively problem-solve partnership barriers may enhance team efficiency and the quality of EBIs/EBPs implementation (9, 10).

At the sustainment/support system level, D&I research requires consideration of the sustainment and sustainability of EBIs/EBPs. Sustainment is the continued use of EBIs/EBPs within practice settings (12); sustainability is the extent to which the EBIs/EBPs can be delivered with their “intended benefits” over an extended period of time after external support from the donor agency terminates (13). Thus to support sustainment and sustainability, D&I research requires PEC between implementation team members and external partners (e.g., patient advocates, EBI/EBP providers, funders, researchers, institutions, community-based organizations, relevant policymakers, and healthcare system partners) (4, 14). Such cross-disciplinary and cross-organizational partnerships address potential structural and system-level barriers and to the expansion and sustainability of EBIs/EBPs.
(5). Partnership barriers that commonly occur at this level include conflicts between: (1) priorities and competing demands across organizations or communities; (2) leaders’ and partners’ roles; and (3) models of partner/community relationships (15). Therefore, utilizing PEC strategies to address conflicts and promote partner engagement and cross-disciplinary partnerships will have important implications for gaining greater support in implementing and sustaining EBIs/EBPs.

The Study Aims
While partnerships in D&I research commonly occur at multiple levels, there has been no multi-level conceptual model to guide PEC strategy development or testing. PEC research is often carried out separately for different partnership levels, and commonly focuses on one level at a time. To inform the development of an integrated D&I framework for PEC strategies, it is important to understand and summarize current research on each partnership level, especially related to the core components and theoretical processes that contribute to effective PEC strategies. Thus, the overall goal of this paper is to address D&I knowledge gaps by reviewing PEC literature and synthesizing knowledge to guide the development of a multi-level PEC theoretical framework. The review focuses specifically on PEC factors influencing PEC processes and outcomes, theoretical frameworks, and evidence from testing of PEC strategies. Given that the central component of the partnership is interpersonal relationship building, we expected that the literature would identify core components that work across different levels of partnerships. The review was therefore synthesized separately for the three levels of partnership. This paper was not intended to be an exhaustive review of the literature, but rather to provide a high-level view of the approaches in which multi-level PEC strategies are studied in D&I contexts.

METHOD
Definitions
Terms related to PEC have been widely used interchangeably and inconsistently. For the purpose of this study, definitions from an array of review papers, as detailed below, were applied to guide our review (5, 14, 16). Review papers were selected based on the inclusion criteria described in the Method section.

Partnership
In D’Amour et al.’s review paper, partnership is defined as “two or more actors join[ed] in a collaborative undertaking (or a set of common goals and specific outcomes) that is characterized by a collegial like relationship that is authentic and constructive.” A partnership can be a relationship between as few as two partners or it can involve a larger number of individuals from groups and organizations (e.g., a network, coalition, or consortium) (5). Under this definition, partnership research has focused on approaches to developing partnerships (e.g., formalizing, sustaining, and ending partnerships) and strategies to build strong working relationships (e.g., cognitive, emotional, and behavioral strategies) (17).

Engagement
Based on Concannon et al.’s review, engagement is defined as “A bi-directional relationship between the patient (or consumer, family) and provider (or implementer) or between the partner and researcher that results in informed decision-making about the selection, conduct, and use of research or interventions” (14). Under this definition, engagement research has focused on strategies to build strong bi-directional relationships between the partners that enhance trust, commitment to collaborate, shared decision-making, problem-solving, and behavioral changes. The terms engagement and alliance are often used inter-changeably (14).

Collaboration
Based on Mattessich et al.’s review (16), collaboration is defined as a mutually beneficial and well-defined relationship entered into by two or more organizations to achieve common goals. Collaborations include a commitment to mutual objectives, a jointly developed structure, shared responsibility, mutual authority, and accountability for success, and sharing of resources and rewards (16). Under this definition, collaboration in research has focused on strategies to enhance partners’ ability to work together to achieve mutual benefits (17). Collaboration is conceptualized as distinctive from cooperation and coordination, which represent earlier stages of organizational partnership (16). Specifically, cooperation is characterized by informal relationships (that exists without any commonly defined mission or planning effort), informal information sharing, preserved authority in each organization, and separated resources by organizations. Coordination is characterized by a more formal relationship, an understanding of compatible missions, with some planning and division of roles, and some established communication channels (16).

Taken together, based on the listed definitions, partnership can be conceptualized as a broader umbrella term that includes engagement and collaboration. Partnerships can occur in multiple forms and at different levels. Therefore, in our review, we included inter-related PEC literatures and diverse types and forms of partnerships.

Literature Review Methods
A cross-field scoping review was conducted, and the review was carried out separately for three forms of partnerships that are commonly applied in D&I research (described above). The scoping review method was used because it provides a useful initial approach to generate foundational knowledge (for each level of PEC research), and to inform approaches for future systematic review (18). In the scoping review, the 5-step method outlined by Arksey and O’Malley(18) was applied. The 5 steps include: (1) identifying the research question (i.e., factors and processes for three levels of PEC); (2) identifying relevant studies/literature; (3) study selection; (4) charting the data; and (5) collating, summarizing, and reporting results. The overall inclusion criteria of articles for this review included studies that: (1) examined partnership, engagement, and/or collaboration factors, processes, mechanisms, or effectiveness of
strategies; (2) examined diverse types and forms of partnerships; (3) had health implications; and (4) were published in English language, peer reviewed literature, from 2000 to 2017, and in PubMed, Ovid MEDLINE, or by credible federal research institutions (e.g., NIH, AHRQ, CDC). Studies that only characterized or described partnership development approaches, and did not examine factors associated with PEC, discuss theoretical frameworks, or assess partnership outcomes were excluded.

To understand consumer-provider level PEC, the literature about patient-provider partnerships, patient/family-centered care, patient/family engagement research that focuses on factors that influence PEC processes and outcomes, and intervention strategies for promoting consumer-provider relationships and engaging consumers to actively use EBI strategies were reviewed. For delivery-system-level PEC, relevant literature about team collaboration, teamwork, inter-professional collaboration, and teamwork interventions that focused on factors that influence PEC processes and outcomes, and intervention strategies for effective team partnership and teamwork were reviewed (9). To understand sustainment/support system level PEC, the literature about multidisciplinary collaboration, quality improvement collaboration, patient-centered outcome research (PCOR), patient/community participation in research, community-based participatory research (CBPR), and collaborative/team science research that focused on factors that influence PEC processes and outcomes, and intervention strategies for effective collaboration across diverse organizations and disciplines were reviewed. These themes were considered because they included diverse partners from multiple organizations, emphasized equitable partnership building, studied factors related to development and sustainment of collaboration, and considered complexity in collaboration process (14, 19–24).

*Figure 1* shows the multilevel PEC conceptual framework that guided this literature review [adapted from Proctor et al. (25)]. The gray boxes represent a summary of the findings from the content synthesis.

**RESULTS**

Tables 1S–3S in the Supplemental file document the charting of review data in detail for studies included in the three levels of PEC literature. Below, findings for each level of partnership are synthesized. Each section is divided into a description of factors that influence PEC processes and outcomes, frameworks used to study PEC strategies, and studies that tested PEC strategies and impact evidence for PEC strategies. A very large body of PEC related studies met our inclusion and exclusion criteria (see number below for each level of PEC review); therefore, priority was placed on review papers, when available. Reviews that examined features for effective PEC and provided approaches for assessing PEC outcomes were included first. Additional selected articles were added when predictors and processes for effective PEC were not covered in the reviewing articles. *Figure 1* and Table 1 show a summary of key results from the cross-level PEC literature review, based on the included articles.

**Consumer-Provider (or Patient-Implementer) Level Partnership**

More than 5,000 articles related to patient-provider partnership and patient/family engagement were identified, along with several review papers. To avoid redundancy, this review focused on synthesizing findings from eight relevant consumer-provider level PEC review papers (four focused on PEC interventions, four focused on PEC related factors) and two framework papers. In total, the selected 10 studies represented findings from 425 research articles.

**Factors That Influence PEC Processes and Outcomes**

The literature outlined four domains that influence effective consumer-provider partnerships. These include: (a) cognitive domain (e.g., providing knowledge; listening and recognizing patients’ perspectives and experiences; assessing patients’ strengths and needs); (b) affective and interpersonal relationship domain (e.g., developing trust, caring, empathetic, respectful, supportive relationship; partnership alliance; identifying and handling emotional problems); (c) behavioral domain (e.g., shared decision-making, providing support, actions to increase EBI/EBP accessibility, actions for finding and trying out solutions to address problems or increase participation, and reinforcement management or homework assignment to increase positive behaviors) (6, 8, 26–32); and (d) contextual factor domain (e.g., health service environment, system, and resources; social determinants; individual partner characteristics). The contextual factors influence not only patient-provider partnership behaviors and processes, but also subsequent PEC outcomes (e.g., cognitive benefit, satisfaction, intervention engagement) (28, 30).

**PEC Frameworks**

The Mutual Participation Model of Care and Transtheoretical Model of Behavior Change (TTM) model have been applied in studying patient-provider PEC processes. The Mutual Participation Model proposes that patient-provider mutual participation and approximately equal power in the treatment process will increase patients’ sense of self-efficacy, improve self-management of health, and increase active participation in treatment (6). The TTM, also known as the Stages of Change Model, proposes that patients move through five stages of change that represent different levels of readiness to engage in behavior change: pre-contemplation/not ready to change, contemplation/getting ready, preparation/ready, action/making change, and maintenance stages. Based on this model, providers are more likely to successfully engage patients in behavior change if they use communication strategies that match recommendations with patients’ level of readiness to change (33). For patients in the early stages of TTM, providers may focus on cognitive and affective strategies to build buy-in, awareness, and trust to prepare for change. For patients in the later stages of TTM, providers may focus on behavioral management, support, and motivation strategies to build strong relationship that support and maintain patient’s change (33).
PEC Strategies Testing

Based on findings derived from four review articles, which synthesize results from 100 intervention studies, most intervention research has targeted providers and patients separately. In the interventions that targeted providers, most strategies tested were communication and consultation strategies, particularly focused on psychological and relational aspects of communication. Strategies included helping providers gain skills in identifying and managing patients' emotional states, sharing decision-making, demonstrating empathy, and seeing each patient as a whole and unique individual (34, 35). In the interventions that targeted the patients/consumers, most strategies have focused on promoting patients' cognitive preparation (e.g., psychoeducation that promotes knowledge, realistic expectations, and participation in EBIs), increasing use of assessment (i.e., assessing patients' barriers to participate and then discussing solutions with patients), and increasing participation (e.g., promoting access to services, increasing attendance and/or adherence) (26, 36).

Impact Evidence for Individual Level PEC Strategy Testing

Evidence showed that interventions focused on provider communication/consultation style training (with 52 randomized controlled trial [RCT] studies out of 60 studies included in the review papers) resulted in significant impact on improving consultation processes, providers' communication skills, and patient satisfaction (34, 35). However, the effects of such interventions on patient healthcare behaviors and health outcomes were limited. Only complex interventions directed at both providers and patients that included condition-specific educational materials demonstrated greater health benefits for patients compared to single component targeted interventions (34, 35). For interventions focused on patient engagement (with 40 RCT studies included in the selected reviews), researchers found that assessment, strategies that promoted access to services, and psychoeducation were more likely to improve patients' engagement in EBIs (measured by attendance, adherence, and cognitive preparation) compared to interventions that did not use these strategies (26).
<table>
<thead>
<tr>
<th>Frameworks</th>
<th>Consumer-Provider level PEC</th>
<th>Implementation Team level PEC</th>
<th>Sustainment/Support System Level PEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>The mutual participation model:</td>
<td>Effective patient-provider partnership is developed through mutual participation and approximately equal power in treatment process</td>
<td>Team process framework: Teamwork is developed through a process of Multilevel Contexts ➔ Team Processes (team cognitive, interpersonal, affective, motivational, and behavioral) ➔ Team Effectiveness</td>
<td>Team efficiency framework: Multidiscipline PEC is developed through a process of Contexts ➔ Processes ➔ Collaboration Outcomes</td>
</tr>
<tr>
<td>The transtheoretical model of behavioral change (TTM):</td>
<td>Consumer-provider level PEC and patient health behavior change are guided by a staged approach of change</td>
<td>Integrated framework: Effective team PEC considers two domains of behaviors that function to regulate a team’s performance and manage team maintenance (keeping the team together)</td>
<td>Social exchange theory: An individual/agency joins a group for exchange purposes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PCOR framework: Effective PEC considers (a) foundation factors (awareness of PCOR method, value for patients); and (b) engagement and action of PEC factors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CBPR framework: Effective PEC considers contextual factors and group dynamic factors (structural, individual, relational dynamics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stage process framework: Factors for effective PEC depend on the partnership stage</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Trust framework: PEC outcomes hinge largely on trust and 3 related factors (openness, transparency and diversity)</td>
</tr>
<tr>
<td>Team science concept map framework:</td>
<td>Effective partnership requires consideration of team factors, support factors, and meta factors</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Features for effective PEC (Factors influence PEC behaviors, processes, and subsequent PEC outcomes)</th>
<th>Features for effective PEC (Factors influence PEC behaviors, processes, and subsequent PEC outcomes)</th>
<th>Features for effective PEC (Factors influence PEC behaviors, processes, and subsequent PEC outcomes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Multi-Level contextual factors: e.g., service environment, resources, individual partner characteristics</td>
<td>• Multi-level contexts: see contextual domain in consumer-provider level PEC</td>
<td>• Collaboration foundation factors: (i) Collaboration environment (e.g., history, interdependence, flexibility, reflection on process, collective ownership); (ii) Team composition (e.g., team diversity, disciplinary dynamic, team membership); and (iii) Organization characteristics (e.g., resources, staff, time, incentives, skilled leadership)</td>
</tr>
<tr>
<td>• Cognitive factors: e.g., recognizing patients’ perspectives and experiences, assessing patients’ strengths, and needs, providing information using individualized approach</td>
<td>• Team cognitive process factors: e.g., team mental model, team learning, commitment</td>
<td>• Factors by processes of collaboration: (i) Cognitive process factors; (ii) Interpersonal, motivational and affective process factors; and (iii) Behavioral process factors (similar to factors listed under Implementation team-level PEC factors/strategies)</td>
</tr>
<tr>
<td>• Affect and interpersonal relationship factors: developing trust, caring, empathetic, respectful, supportive relationships; building a partnership alliance; identifying and handling emotional problems</td>
<td>• Team interpersonal, motivation, affective process factors: e.g., team cohesion, team efficacy, team affect, team conflict, psychological safety</td>
<td>• Factors by stage of collaboration: (i) During the preparation stage may include interpersonal and operational related factors (e.g., sharing goals, setting the stage for an engaged and supportive organizational culture, developing institutional structures to address and support potential system barriers, mutual respect, building partners’ knowledge and skills); and (ii) during the execution stage may include partnership synergy, knowledge exchange, monitoring, and ongoing support-related factors (e.g., co-learning strategies, building reciprocal, equal, genuine, and trust-worthy relationships, re-assessment and feedback)</td>
</tr>
<tr>
<td>• Behavioral factors: sharing control of the decisions and responsibility; providing support; taking action to increase EB/EBP accessibility, and for finding and trying out solutions to address problems or increase participation; and reinforcement management</td>
<td>• Team behavioral process factors: (i) factors related to regulating team performance (e.g., goal-specific planning before a team task; task-related collaboration and team monitoring during team task execution; and team practice innovation after team task); (ii) factors related to team management (e.g., support and conflict management)</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
During QIC set-up, 7 key PEC factors have been targeted and tested (i.e., pre-work-convened expert panel, PEC strategies tested in quality improvement collaboration (QIC) pre-work-organization commitment, in-person learning sessions, PEC Team member training strategies). Provider training strategies have focused on team regulation (strategies to keeping the team together during execution period email or web support). Teamwork preparation, execution, and external support with monitoring data synthesis have been applied. At the organization level, leadership involvement and interpersonal dynamics, such as conflict management and psychological support strategies, have been tested. At the individual level, strategies such as providing CBPR training (to build partners capacity for partnership research), developing structured communication mechanisms to facilitate opportunities for discussion and partnership skills building, and providing technical assistance on research related design have been tested.

Factors That Influence PEC Process and Outcomes and PEC Frameworks

Factors that influence PEC processes and outcomes were based on two representative teamwork frameworks (10, 37, 38). Specifically, the Integrated Framework for effective teamwork, developed by Rousseau et al. based on a review of 29 studies that examined teamwork behaviors and processes, proposes that teamwork behaviors are constructed in a nested hierarchical structure (10). Effective team PEC needs to consider two domains: (1) behaviors that function to regulate a team’s performance; and (2) management of team maintenance. With regard to regulating team performance, PEC strategies need to include those that occur (a) before team task performance (e.g., creating action plans, team mission analysis, goal setting, cognitive preparation for team task performance); (b) during the execution of team performance (e.g., coordination, cooperation, and information exchange, task-related collaborative behaviors/strategies, monitoring team performance, reflection); and (c) after task team adjustment period (e.g., intra-team coaching, collaborative problem-solving, and team practice innovation). With regard to management of team maintenance, PEC strategies may include psychological support and integrative conflict management (10).

Different from Rousseau’s framework (10), which focuses more on PEC strategies based on the stage of team development, Kozlowski et al. (37, 38) proposed a Team Process Framework that posits that the context in which a team works influences team processes, which in turn influence team effectiveness and performance. In Kozlowski’s model, team PEC needs to be conceptualized in a multilevel context (considering individual, organizational system, and environmental influences). Moreover, in order to have effective team-level PEC, partnership factors in three distinct but inter-related team processes need to be considered, including: (a) cognitive team processes (e.g., collective team climate and safety climate, team mental models, team learning factors); (b) team interpersonal, motivation, and affective processes (e.g., team cohesion, team efficacy, team affect/emotion/conflict); and (c) team action and behavioral processes (e.g., team coordination/cooperation/communication, team competencies/functions, team regulation, performance dynamics, adaptation) (37, 38).

Authors of other review and theoretical perspective papers (39, 40) also suggest that promoting teamwork requires
similar processes to the frameworks described above [e.g., in a review paper of teamwork monitoring instruments, most have focused on team contexts and behavioral processes based on the two conceptual frameworks described above (40)]. In team contexts, team composition and structure, organizational climate, individual attitudes, beliefs, values, and commitment about teamwork were commonly assessed. In assessing team behaviors, collaborative behaviors, such as communication, goal settings, task analysis, monitoring, adjustment collaboration, problem-solving, decision-making, workload sharing, conflict, and team leadership were commonly assessed (40). Team climate, including climate related to psychological safety, team objectives, team commitment, and support for innovation, has also been proposed for fostering effective team PEC and recommended for carefully monitoring (39). These commonly assessed constructs represent the importance of these factors in team PEC processes.

**PEC Strategies Testing**

In a meta-analysis based on 72 interventions from diverse fields, researchers reported that most intervention strategies have targeted team member training and most content designs were based on the Integrated Framework (10). Strategies commonly included were related to team regulation strategies (e.g., strategies to keep teams engaged during teamwork preparation, execution, and reflection) and team maintenance strategies (e.g., conflict management and psychological support) (9). Other training models applied holistic/humanities and team co-learning approach. This training approach focused on the patient holistic care concept and provided tools and opportunities to facilitate team members’ co-learning and inter-professional team collaboration to provide holistic patient care (41).

**Impact Evidence for Team-Level PEC Strategy Testing**

Authors of the meta-analysis reported that overall, team training had significant, medium-sized effects in enhancing both teamwork and team performance across a variety of team contexts and training methods (9). In addition, regardless of the targeted domains (e.g., preparation, execution, reflection, interpersonal dynamics) and number of teamwork domains targeted, teamwork training significantly improved team performance. However, in terms of improving teamwork behaviors, significant effects only emerged when two or more domains of teamwork were targeted (9). Trainings using the holistic/humanities and team co-learning approach resulted in significant improvements in team efficiency, team value, shared roles, knowledge, satisfaction, and reactions to working in team across all levels of learners (including non-English-speaking and diverse provider staff) (41).

**Sustainment/Support System Level PEC**

To understand sustainment/support system-level partnerships, literature on interdisciplinary collaboration, quality improvement collaboration, patient/community research partnership research (e.g., PCOR, CBPR, patient/community participation in research), and team/collaborative science was reviewed. More than 7,000 articles were identified. For this review, 22 relevant studies were included, 9 of which were review papers and 13 of which were frameworks or empirical studies that examined PEC factors or/and processes. The reviewed 22 studies represented findings from 597 research articles.

**Factors That Influence PEC Process and Outcomes**

Factors and processes for two key topic areas of literature—interdisciplinary collaboration and patient/community-academic partnership research—were examined separately given the rich and diverse topics within each field of research. Twelve studies that described interdisciplinary collaboration research were reviewed in detail. These included studies on interdisciplinary, quality improvement collaboration [QIC], and team/collaborative science (six were reviews). Overall, results revealed that factors influencing effective PEC mapped onto two broad domains: (a) Factors related to team foundation; and (b) factors related to processes. Factors frequently studied under team foundation were related to collaboration environment (e.g., history, political/social climate, interdependence, flexibility, reflection on process, collective ownership, mutual respect, ability to compromise, trust), team composition (e.g., team diversity, disciplinary dynamic, multiple layer of participation, representation of organization), and organization characteristics (e.g., resource, fund, staff, time, incentive, skilled leadership). Better understanding of team foundation factors and PEC contexts can guide the use of PEC strategies to prepare PEC set-up and increase partners’ readiness for PEC. Factors frequently studied under the processes were related to cognitive processes (e.g., clear roles, shared visions/values, concrete attainable goals, cross-disciplinary learning), interpersonal/motivational/affective processes (e.g., established informal relationships, communication mechanisms, value the contribution of collaborators), and behavioral processes (e.g., having open and honest communication, sharing decision-making, power sharing, acknowledging egalitarian nature of relationships, identifying barriers, and problem-solving) (5, 16, 42–47).

Separate from the interdisciplinary collaboration literature, eight studies on patient/community-academic partnership research, including literatures from PCOR, CBPR, and patient/community research partnership research, were also reviewed (four were reviewed studies). Overall, similar PEC factors were identified as in the interdisciplinary collaboration literature, as well as in the implementation team-level partnership literature (described above). However, there were some differences in two areas of research. Patient/community-academic partnership literature was more likely to discuss factors or strategies based on stages of partnership (rather than domains). Factors or strategies frequently studied during the preparation period were interpersonal and operational process related strategies. These might include sharing goals, establishing an engaged and supportive organizational culture, developing institutional structure to address and support potential system barriers, developing mutual respect, and building partners’ capacity for partnering skills (22, 24, 48, 49). Factors or strategies frequently studied during the PEC execution period were...
partnership synergy, knowledge exchange, monitoring, and support related strategies (e.g., co-learning strategies, building reciprocal/equal relationships, assessment, and feedback) (22, 24, 50).

PEC Frameworks
There are several conceptual frameworks (49) for explaining the effect of PEC at the level of sustainability and support systems. Because of the complexity of partnership at this level, more PEC frameworks have been developed. Conceptual frameworks, such as the Team Efficiency Framework, Social Exchange Theory, PCOR, CBPR, Stage Process Framework, have been applied in interdisciplinary/interagency PEC research (5, 51–54). The Team Efficiency Framework, which is commonly applied in team member partnership (described above), proposes that multi-disciplinary collaboration is a process/configuration of input (contextual factors) process (cognitive, relationship/affective, and behavioral processes) outcomes (performance, innovation, viability) (55). The Social Exchange Theory proposes that an individual/organization joins a group for exchange purposes. The partnership provides specific benefits to individuals/organizations and that, in return, the individuals/organizations are expected to help the group attain its objectives. From this perspective, the challenges that commonly occur during partnerships are related to concerns about power-sharing in attaining equal benefits (56). Therefore, PEC strategies focused on power, decision-making, and interaction dynamics are commonly proposed. The PCOR Framework, developed by the Patient-Centered Outcomes Research Institute/PCORI, emphasizes trust, honesty, co-learning, transparency, reciprocal relationships, partnership, and respect in collaboration processes. PCOR proposes two broader domains of PEC factors to be considered: (a) contextual factors: including internal factors such as awareness of methods for PCOR, a patient centered culture, and external factors (e.g., ways for patients and researchers to partner, resources and infrastructure, policies and governance); and (b) engagement action of PEC factors: including initiating and maintaining partnership; facilitating cross-communication among partners; capturing and optimizing partners’ perspective across phases of research; ensuring meaningful influence on research; providing training for partnering; and sharing and applying learnings (52).

The CBPR Framework is also based on trust, respect, mutual benefit, and equitable and shared decision-making principles similar to the PCOR framework (57, 58), and proposes two overarching domains that are relevant to PEC (53, 54). The contextual domain includes contextual factors that influence partnerships, including: social, economic, cultural, local/national governance, policies, and funding trends, role of institutions, historical context of trust/mistrust, both university and community partners’ capacities, readiness, and experience in participatory research, and perceived severity of health issue. The group dynamic domain considers three areas of factors that influence PEC dynamics. These include: (a) structural dynamic (e.g., diversity, complexity, formal agreements, real power/resource sharing, alignment with CBPR principles); (b) individual dynamics (e.g., core values, motivations for participating); and (c) relational dynamics (e.g., safety, trust, flexibility in dialogue, listening and mutual learning, leadership influence, power dynamics, self, and collective reflection, participatory decision-making) (54). It is conceptualized that positive collaboration contexts and group dynamics will result in positive synergistic partnerships, appropriate interventions, and research, and improved systems and community capacity (53, 54).

The Stage Process Engagement Framework (proposed by NIH, CDC and other researchers) suggests that key factors for effective PEC depend on the stage of collaboration (51, 59–61). At the initial stage, PEC may focus on clarifying collaboration goals, promoting knowledge about the collaborators, and better understanding the strengths and weaknesses of the partnering contexts. For engagement to occur, it is necessary to visit communities to establish relationship and build trust, and subsequently work toward developing mutual beneficial goals. For engagement to succeed, sharing responsibility, recognizing and respecting the diversity of partners/communities, and creating transparency are necessary. For partnerships to be sustained, mobilizing community assets and strengths, developing the community capacity, resources, and social capital to facilitate creation of innovative strategies, releasing control of action to the community, and being flexible enough to meet changing needs are also needed (51, 59).

Other frameworks derived from team science literature can also be applied to study multi-disciplinary collaboration process. The Trust Framework proposes that successful PEC outcomes hinge largely on the most basic of human relationship “trust.” The nature of complex collaborative relationships is shaped and formed by three trust related factors: openness, transparency, and diversity. High levels of openness (in team social network) and transparency (related to information and knowledge sharing) will foster diversity in innovation. All three factors are required and need to be balanced for the eventual win-win-win success (62). The Team Science Concept Map proposes not only team related factors, but also support and meta factors that influence the performance of PEC need to be considered. Examples for the team factors may include disciplinary dynamic; structure and context for team; and characteristics and dynamics of teams. Examples for the support factors may include institutional support and professional development and management and organization for team. Examples for the meta factors may include definitions of team collaboration and models, measurement, monitoring, and evaluation (63).

PEC Strategies Testing
Some strategies focused on the sustainment/support systems level were identified from the literature on Quality Improvement Collaboration (QIC), and PCOR and CBPR patient/community research partnerships. In QIC research, several PEC strategies have been tested by researchers during QIC set-up and execution periods. Specifically, during QIC set-up, 7 key PEC strategies were commonly tested. These include pre-work-convened expert panel, pre-work-organizational commitment, in-person learning sessions, Plan-Do-Study-Act/ PDSAs cycles, multidisciplinary team, team calls, email, and/or web support). During the
execution period, strategies such as monitoring data collection, reviewing data for feedback, and using external support for monitoring data synthesis and feedback have also been applied. At the organizational level, PEC strategies such as involving leadership and providing QIC training for staff members were applied (19). In PCOR research, PEC strategies developed from the PCOR engagement framework were also studied in PCORI funded projects (52). In CBPR research, training strategies based on CBPR framework have also been studied. The goals of these CBPR trainings are to build partners capacity, develop structured communication mechanisms to facilitate opportunities for discussion, develop partners’ partnership skills, and capacity, and provide technical assistance on research related design (48).

PEC Evidence at the Sustainment/Support System Level
Research in this level of PEC strategy testing is limited and relies more on qualitative and short term data collection. A QIC review study (based on 24 RCTs or quasi-experimental studies) found some positive evidence for PEC strategies. In general, the impact of QIC tends to be greater for providers than for patients. At the provider level, about 47% QIC studies showed positive findings (42% mixed findings and 11% no findings) related to patient-centered cares, such as showing improvement on patient health screening/monitoring, use of data to inform interventions, and/or provider teamwork. At the patient-level, only 23% studies showed positive findings (46% mixed findings, and 31% no findings) related to an increase in patients’ participation in care or reduction in health symptoms (19). Regardless, findings were not surprising because most QIC focused on provider related PEC strategies.

For PEC strategy testing based on the PCOR framework, findings from a recent study of 221 PCORI funded projects between 2012 and 2016 (based on self-report data from 235 investigators and 260 partners) provide some supporting evidence for the PCOR approach of collaboration. There were 11–52% investigators and partners endorsing improvements on patient-centeredness of study processes and outcomes (e.g., choices of research topics were driven by patients and related to their needs), and 20–81% investigators and partners endorsing improvement in study design, conduct, or/and efficiency (e.g., increasing the appropriateness of research question selection, design, and outcome measures) (52).

Other PEC strategy testing studies have been based on the CBPR framework and have used provider training and technical support to researchers and community partners to promote effective community-academia collaboration. Positive findings have been documented in several studies, especially related to achieving deliverables (e.g., written pilot study proposal, IRB approved study protocol, carried out pilot studies) (48, 64–66).

DISCUSSION
The purpose of this paper is to describe a PEC framework and methodological gaps in D&I research by reviewing and summarizing findings from a broad range of PEC literature. A total of 39 articles (including 21 review articles) were selected for this review, representing findings from 1,456 research studies. Through this review, factors and theoretical processes that influence PEC, and strategies that promote effective PEC were identified. Findings guided the development of a multi-level PEC framework, which can be applied to strengthen the evidence-based for PEC research in the field of D&I.

In identifying factors for effective PEC, four domains were consistently identified across three levels of PEC. These included cognitive, interpersonal/affective, behavioral, and context preparation (see Table 1 and Figure 1 for summary) (28, 30, 37, 38). Furthermore, certain strategies were found to be more critical based on the partnership stage. In the earlier stages of PEC, cognitive, affective, and experiential behavioral change strategies were more important in order to build buy-in, awareness, mutual goal development, and trust to prepare for partnership (33, 37, 38, 51, 59). In regulating partnership performance, cognitive and behavioral strategies were more relevant, and in management of partnership maintenance (keeping the partnership going), relationship and affective strategies were more relevant (10). In inter-organizational types of partnerships, structural dynamics, and related strategies (e.g., alignment of collaboration goal with agency mission, resource sharing, leadership representation, and power) need to be carefully considered because of their potential influence on relational dynamics (e.g., integration of agency beliefs to team partnership process) (54).

Related to PEC mechanisms, several useful theoretical frameworks that include factors related to PEC processes were identified. For example, the relationship among context-mechanisms-outcome (or input→process→outcome) has been used to develop causative explanations about PEC processes. This approach allows process modeling wherein the outcome of one context-mechanism-outcome becomes the context for the next chain of implementation steps. Although this framework is useful to guide PEC process research, it may not be as useful for studying PEC mechanisms at the sustainment/support system level due to the complexity of behavioral dynamics within and across agencies during different partnership stages. Recently developed integrated frameworks of change, such as the stage of change (e.g., TTM), partnership stage of change (e.g., cooperation-coordination-collaboration; preparation-execution-adjustment) (33, 37, 38, 51, 59, 60), and multiple dynamic processes (e.g., cognitive, interpersonal relationship, behavioral dynamics) (37, 38, 53, 54) may generate more complex and explanatory theories to guide the design of PEC strategies in D&I research.

Leadership is a factor that is frequently studied in D&I contexts. In this review, we found that leadership’s function varies based on stages or levels of partnership. At the delivery system level (or implementation team level), team leadership plays a role during the task execution period (to facilitate activity coordination) (40). At the sustainment-level of PEC, skilled/effective leadership is considered to be a collaboration foundation strategy, which plays a supporting role in the
interdisciplinary collaboration process and facilitate initial institutional structure set-up to support PEC (15, 16, 51). In inter-organizational collaborations, leadership also plays an important role given the involvement of different partnering organizations (e.g., in negotiation of collaboration goals), and the importance of facilitating cross-communication among agency staff and external partners (5, 19). Therefore, leadership strategies should be considered when they are relevant to study design.

In PEC strategies testing, four approaches were found to be effective. Training (or psychoeducation) for consumers, providers/implementers, and partnering members was consistently identified as an important PEC strategy across all level of partnerships. Training provides an opportunity for cognitive preparation and skill building to allow partners to communicate more effectively and to actively participate in partnership activities. Training can also target positive relationship building, partnership behavior engagement, and partnership sustainment (9, 26, 35, 41, 48). However, training alone does not change health behaviors. Training that incorporates multiple strategies and targets all partners is more likely to change behavior (9) and to improve patient health benefits (3, 35).

In addition, strategies that focus on assessment/monitoring/reflection (e.g., partner members’ ability to assess barriers or collect monitoring data for feedback), participation (e.g., power sharing, involvement in decision-making), and relationship building (e.g., communication style, conflict management) are useful and could be included in training initiatives (9, 19, 26, 34, 35).

Implications for D&I Research
Three main lessons for D&I research can be drawn from this review. First, researchers may want to consider gathering data about PEC contexts, associated factors, and processes at multiple levels as part of initial assessments of D&I contexts to enable examination of how PEC contexts and processes from each level of partnership contribute to service use, patient health, and sustainment outcomes [as defined by (25)].

Second, several research questions emerged. Most tests of PEC strategies have only evaluated short-term or intermediate outcomes. Limited evidence is available related to long-term effects. In addition, most sustainment/support PEC strategies have only been evaluated as case studies, qualitatively, or in non-experimental designs. Integrated conceptual frameworks have only recently been developed that could elucidate the complexities of analysis of sustainment strategies. Furthermore, the lack of measurement tools for construct assessment has been an impediment. As measurement tools become more refined and feasible, future studies can take advantage of these advances to study these types of strategies.

Third, PEC has not yet been integrated into D&I training. While partnerships are common, true power sharing is rare. Many D&I experts recognize the importance of including PEC strategies, but do not systematically incorporate them into training and evaluate impacts on PEC effectiveness. It would be useful to include training in specific skills related to PEC. These might include team science, community engagement, communication strategies, conflict management, and interpersonal and intrapersonal intelligence (67, 68).

CONCLUSION
As population level health issues continue to require complex healthcare policy solutions, it will become increasingly important to improve partnerships, engagement of different constituents, and new collaborations to craft cost-effective and creative solutions on a broad scale. This will entail at a minimum active involvement of patients, policy-makers, providers, community leaders, and researchers. This paper provides several new directions to address D&I knowledge and methodological gaps related to these partnerships. The review and framework not only provides guidance on how PEC related factors and outcomes can be conceptualized, but also how PEC processes can be integrated into more robust D&I designs. As has been reiterated, more research is needed to elucidate both cross and multi-level partnership mechanisms. In particular, systematic and long-term follow-up research will strengthen understanding of PEC strategies to advance EBI/EBPs implementation-effectiveness, sustainability, and system and population health outcomes.

AUTHOR CONTRIBUTIONS
K-YH, SK, LB, DS, SAK, OO, and KH contributed conception and design of the study. K-YH, SK, and SC was involved in the acquisition, analysis, and interpretation of data. K-YH wrote the first draft of the manuscript. K-YH, SK, SC, DS, DK, and KH contributed to manuscript writing. All authors contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS
This research was supported by grant NIH/NCATS U19 MH110001-01 from the National Institute of Mental Health (NIMH). The reviews and opinions expressed in this article are solely those of the authors and do not necessarily represent the views of NIH. This review relied on publicly available documents and, therefore, was exempt from Institutional Review Board determination.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2018.00190/full#supplementary-material
REFERENCES


27. Brennan SE, Bosch M, Buchan H, Green SE. Measuring team factors thought to influence the success of quality improvement in primary


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Huang, Kwon, Cheng, Kamboukos, Shelley, Brotman, Kaplan, Olgubenga and Hoagwood. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Variation in Research Designs Used to Test the Effectiveness of Dissemination and Implementation Strategies: A Review

Stephanie Mazzucca1*, Rachel G. Tabak1, Meagan Pilar1, Alex T. Ramsey2, Ana A. Baumann3, Emily Kryzer3, Ericka M. Lewis4, Margaret Padek1, Byron J. Powell5 and Ross C. Brownson1,6

1 Prevention Research Center in St. Louis, Brown School, Washington University in St. Louis, St. Louis, MO, United States, 2 Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, United States, 3 Brown School of Social Work, Washington University in St. Louis, St. Louis, MO, United States, 4 School of Social Work, University of Maryland, Baltimore, MD, United States, 5 Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, 6 Department of Surgery, Alvin J. Siteman Cancer Center, Washington University School of Medicine, Washington University in St. Louis, St. Louis, MO, United States

Background: The need for optimal study designs in dissemination and implementation (D&I) research is increasingly recognized. Despite the wide range of study designs available for D&I research, we lack understanding of the types of designs and methodologies that are routinely used in the field. This review assesses the designs and methodologies in recently proposed D&I studies and provides resources to guide design decisions.

Methods: We reviewed 404 study protocols published in the journal Implementation Science from 2/2006 to 9/2017. Eligible studies tested the efficacy or effectiveness of D&I strategies (i.e., not effectiveness of the underlying clinical or public health intervention); had a comparison by group and/or time; and used ≥1 quantitative measure. Several design elements were extracted: design category (e.g., randomized); design type [e.g., cluster randomized controlled trial (RCT)]; data type (e.g., quantitative); D&I theoretical framework; levels of treatment assignment, intervention, and measurement; and country in which the research was conducted. Each protocol was double-coded, and discrepancies were resolved through discussion.

Results: Of the 404 protocols reviewed, 212 (52%) studies tested one or more implementation strategy across 208 manuscripts, therefore meeting inclusion criteria. Of the included studies, 77% utilized randomized designs, primarily cluster RCTs. The use of alternative designs (e.g., stepped wedge) increased over time. Fewer studies were quasi-experimental (17%) or observational (6%). Many study design categories (e.g., controlled pre–post, matched pair cluster design) were represented by only one or two studies. Most articles proposed quantitative and qualitative methods (61%), with the remaining 39% proposing only quantitative. Half of protocols (52%) reported using a theoretical framework to guide the study. The four most frequently reported frameworks were Consolidated Framework for Implementing Research and RE-AIM (n = 16 each), followed by Promoting Action on Research Implementation in Health Services and Theoretical Domains Framework (n = 12 each).
BACKGROUND

Dissemination and implementation (D&I) research is a relatively new scientific field that seeks to understand the scale up, spread, and sustainability of evidence-based interventions (EBIs) and practices for broad population health impact. D&I studies focus on effective strategies to enhance the speed of intervention implementation, quality of intervention delivery, and the extent to which the intervention reaches those it is intended to serve (1–4). D&I research is the final stage of the research to practice pipeline, and several characteristics of D&I studies differentiate them from efficacy and effectiveness studies. The exposures (the independent variables) in D&I studies are D&I strategies, whereas in efficacy and effectiveness studies, the exposures are the EBIs themselves (4). In D&I studies, outcomes are often related to the speed, quality, or reach of intervention implementation or delivery; these are often proximal outcomes, processes, and outputs of the service delivery system, and sometimes distal patient-level outcomes (1–4). As such, D&I studies are inherently multilevel, and accurate evaluation requires an understanding of the levels at which interventions are tested, implemented, and measured (5). D&I study outcomes are distinct from those in efficacy and effectiveness trials, which are related to changes in the target behaviors of end users or determinants of those behaviors (3). Due to the differences in D&I studies compared to efficacy and effectiveness studies of underlying interventions, the prioritization of study design considerations and study designs needed for D&I research are likely different than those of efficacy and effectiveness studies.

Traditional study designs such as randomized controlled trials (RCTs) can be ideal for testing the efficacy or effectiveness of interventions, given the ability to maximize internal validity. However, there has been concern that traditional designs may be ill-suited for D&I research, which requires a greater focus on (a) external validity; (b) implementation-related barriers and facilitators to routine use and sustainability of “effective” practices (6); (c) studying factors that lead to uptake of effective practices at the organizational level; and (d) capturing “moderating factors that limit robustness across settings, populations, and intervention staff,” including race/ethnicity, implementation setting, or geographic setting (7). Designs that enhance external validity allow us to better understand how interventions and implementation strategies work under realistic conditions rather than in highly controlled circumstances.

A number of alternative designs are available that give researchers flexibility and allow them to maximize external validity, match the research question of interest appropriately with the phase of D&I research (i.e., exploration, preparation, implementation, and sustainability) (4, 8, 9), and balance other trade-offs influencing the choice of design (10) (e.g., if randomization is appropriate, preference of stakeholders, etc.). If a randomized design is desired, it may be necessary to consider non-traditional ways to randomize, such as by time, to balance internal and external validity (4), and the practical, ethical, and pragmatic considerations that make some randomized designs less appealing in D&I research (4, 6, 9). For example, there is an ethical justification for designs that allow all stakeholders to receive an EBI and/or D&I strategy that is thought to be efficacious (11), since D&I studies focus on changes in organizations and communities led by stakeholders in these settings who often have more at stake than researchers (9). If a randomized design is not appropriate, other design features can be used to increase internal validity, such as multiple data collection points before and after the EBI is implemented (9). The evaluation of D&I strategies focuses on the process of implementation and stakeholders’ perceptions of this process (12, 13), and the choice of study design depends in part on the preferences of these stakeholders. Thus, a variety of designs that accommodate these considerations will likely be necessary to respond to calls from the National Academy of Medicine (formerly the Institute of Medicine) and numerous other organizations to accelerate the reach of EBIs and close gaps in the quality of health care and public health efforts (14–20).

Some of the alternative designs that are particularly suited to D&I research include interrupted time series, factorial designs, and rollout designs. An interrupted time series (21), in which multiple observations are taken before and after implementing an EBI, might be ideal when selecting the most cost-effective EBI and implementation strategy in the exploration phase. A factorial design, in which the combination of multiple D&I strategies are tested, could be more useful when testing the effectiveness of several different implementation strategies alone or in combination in the implementation phase. Adaptive designs are those in which study characteristics (e.g., implementation strategy type or mode) change throughout the study and may be useful when determining the sequence and combination of implementation strategies (22). Additionally, rollout designs (9), in which the timing of EBI implementation is randomly assigned, are a broad category of designs that include stepped wedge designs (23), where sites continue with usual practice until randomly assigned to transition to the EBI implementation for a defined period. These rollout designs may be more appealing or seen as more ethical to stakeholders than a cluster randomized trial with a no treatment control group, since all participants receive the D&I strategy and intervention packages at some point during the study period (24). There are many considerations that contribute

Conclusion: While several novel designs for D&I research have been proposed (e.g., stepped wedge, adaptive designs), the majority of the studies in our sample employed RCT designs. Alternative study designs are increasing in use but may be underutilized for a variety of reasons, including preference of funders or lack of awareness of these designs. Promisingly, the prevalent use of quantitative and qualitative methods together reflects methodological innovation in newer D&I research.

Keywords: research study design, research methods, review, implementation research, dissemination research
to the choice of design, and assessment of the designs currently being used in D&I research is needed so that future implementation efforts may better account for these differences as well as the contextual factors and multiple levels involved in this field of study (25, 26).

This review was inspired by workgroup meetings supported by the United States (US) National Institutes of Health, “Advancing the Science of Dissemination and Implementation,” which focused on research designs for D&I research. The workgroup described 27 available designs (27), which have been categorized by Brown and colleagues into three types: within-site designs; between-site designs; and within- and between-site designs (9). Despite the increasing recognition of the need for optimal study designs in D&I research (4, 6), we lack data on the types of research designs and methodologies that are routinely used in D&I research. Therefore, we aimed to fill this gap by exploring the range of designs and methodologies used in recently proposed D&I studies testing implementation strategies. Our goals were to assess variation in designs and methodologies used, potentially categorize innovative design approaches, and identify gaps upon which future studies can build.

**METHODS**

Study protocols published in *Implementation Science* from 2/22/2006 to 9/7/2017 (*n* = 400 manuscripts) were screened for eligibility (Figure 1). Manuscripts reporting study protocols typically provide detailed information about the study design and levels of intervention implementation and measurement; as such, this review included only study protocols to assess these factors across studies. To identify studies that were likely to use a variety of innovative methods, our search focused on *Implementation Science*, one of the top journals dedicated to publishing D&I research (28) that also has a specific designation for protocols. In addition, the journal has a focus on publishing “articles that present novel methods (particularly those have a theoretical basis) for studying implementation processes and interventions” (29).

Two of the included protocol manuscripts provided the descriptions of three studies each, resulting in 404 studies reviewed. To be included for full review, studies needed to test the efficacy or effectiveness of D&I strategies using some sort of comparison design. Studies were excluded if they were not testing a D&I strategy, if they were only testing the efficacy or effectiveness of a clinical or public health intervention itself, if they were purely qualitative, or if they did not include a comparison involving the D&I strategy (e.g., by group or time). D&I strategies are processes and activities used to communicate information about interventions and to integrate them into usual care and community settings (4, 27, 30–33). We used previous work by Powell and colleagues to categorize implementation strategies (27) to represent both D&I strategies within this review, since there has been more work done to articulate and categorize implementation strategies compared with dissemination strategies and there is likely a high amount of overlap between the strategies for each category of research (34).

A data extraction template was used to code the following design elements: design category (e.g., randomized, observational); design type (e.g., cluster RCT, pre–post no control); data collection with quantitative only or a combination of quantitative and qualitative methods; conceptual/theoretical framework used; levels of assignment, intervention, and measurement (30, 35); and country in which the research was conducted. Reviewers coded design types exactly as they were presented by study authors to capture the variety of terms used for study designs; for example, the same design was referred to as “interrupted time series with no controls” and “pre–post, interrupted time series” in different studies. Hybrid designs, those blending elements of effectiveness and implementation studies in one trial (6), were not specifically coded so that manuscripts published before this term was introduced could be included. Studies that were labeled as a hybrid study by authors were coded according to the design by which authors tested the implementation strategy. Levels of assignment, intervention, or measurement were coded as individual client or provider; groups/teams of clients or providers (e.g., a surgical unit within a hospital); organization (e.g., local health department); or larger system environment (e.g., province) (35). Each protocol was double-coded, and the few discrepancies were resolved through discussion with the study team.

Some have suggested that it is most appropriate to assign to a treatment arm and measure at the level of implementation (i.e., at the level where the full impact of the strategy is designed to occur) (9, 36). Therefore, studies were grouped according to the extent to which there was consistency between design components: the levels of assignment, intervention, and measurement (Figure 2). *No consistency* occurred when design components were all at different levels. *Partial consistency* occurred when there was at least one level with two matching components, but none with three matching components. *Single-level consistency* occurred when intervention components and measurement were at the level of
assignment. *Multilevel consistency* occurred when intervention components and measurement were at the level of assignment and there was at least one additional level with matching intervention components and measurement.

**RESULTS**

**Study Designs**

Of the 404 studies screened, 212 (52%) tested one or more implementation strategy (Figure 1). The most common reasons for exclusion were the studies that did not test an implementation strategy (n = 94, 49%), were an exploratory study (n = 26, 14%), or the studies that did not have a comparison (n = 23, 12%). Of the included studies, 164 (77%) utilized randomized designs, primarily cluster randomized trials (n = 103, 49%), RCTs (n = 28, 13%), or stepped wedge cluster randomized trials (n = 16, 8%, Table 1). Only 35 studies (17%) were quasi-experimental and fewer (n = 13, 6%) were observational. One paper (37) that reported three studies included in this review contained very little information in the manuscript on study design; these studies were determined to be randomized trials according to context provided in the paper and group consensus. There was considerable variation in the way authors described their study designs. For example, “pre–post with controls” and “cluster controlled pre–post” both referred to the same methodological approach. These subtle differences in study design are likely important and reflect differences in the population, data type, and contextual influences available to the study authors. Complete coding for each study is available (Data Sheet S1 in Supplementary Material).

There was a notable increase in the use of alternative designs over time. For example, stepped wedge designs were not used before 2011, but were proposed in at least four studies per year in 2014–2016. Conversely, there was a decrease in the reliance on individual-level RCTs. Between 2006 and 2012, RCTs represented 20% of all studies, whereas they only represented 8% of studies between 2013 and 2017. Additionally, researchers are utilizing a wider range of designs. From 2006 to 2012, there was an average of four types of designs used per year, which increased to 8.8 per year between 2013 and 2017.

**Levels of Assignment, Intervention, and Measurement**

**Assignment**

For most studies (n = 124, 67%), the intervention was assigned at the level of the organization. Twenty-three studies (12%) used assignment at the level of the individual provider, and the remainder of the studies (n = 39, 21%) reported some combination of individual client, individual provider, group/team provider, and organization.

**Intervention**

Interventions were most commonly targeted at the individual provider (n = 51, 27%); the individual provider and the organization (n = 29, 16%); the organization alone (n = 23, 12%); or both the individual provider and client (n = 20, 11%). There were several studies that targeted clients, providers, and the organization (n = 14, 8%); individual providers and groups/teams of providers (n = 14, 8%), or groups/teams of providers (n = 11, 6%). The remaining studies targeted a variety of levels, for example, clients and larger system environments.

**Measurement**

Studies most frequently (n = 45, 24%) measured outcomes at the individual provider and client levels with fewer studies measuring at the level of the client, provider, and organization (n = 32, 17%) or clients alone (n = 21, 11%). Several studies also conducted

---

**Figure 2** Consistency across levels of assignment, intervention, and measurement. Patterns of consistency across study design components are illustrated with eight example studies. Design components included are assignment (i.e., random or non-random allocation to study arms), intervention and/or implementation efforts, and measurement. Studies were grouped by patterns of consistency of levels across design components. The number and proportion of reviewed studies that fall into each consistency pattern are included. *Symbols indicate the presence of a design component at a given level. Levels are defined as:* ● Organization, e.g., hospital, school. ▲ Provider, e.g., doctor, teacher. ■ Client, e.g., patient, student.
The consistency between levels of intervention and measurement was more variable. Most studies had one level of intervention \((n = 56, 30\%)\) or multiple levels of intervention \((n = 55, 30\%)\), which had corresponding levels of measurement. Thirty-five studies \((19\%)\) had some overlap between intervention and measurement levels, for example, studies that intervened at the individual provider and organizational level, but measured at the individual client and provider levels. Forty studies \((22\%)\) had no consistency between intervention and measurement levels, for example, studies that intervened at the provider level, but measured at the client level. Comparing across all three levels, 44 \((24\%)\) studies had multilevel consistency between the level of assignment, intervention, and measurement, while 43 \((23\%)\) were consistent across a single level (Figure 2). Ninety-one studies \((49\%)\) were partially consistent, for example, assignment occurred at the level of the individual provider, intervention occurred at the level of the individual provider, and measures were taken at the level of the individual client.

### D&I Models, Theories, and Frameworks

Included protocols utilized a wide range of D&I conceptual frameworks. One hundred and eleven \((52\%)\) of the studies reported using a D&I model, and there were a variety of models used. The Consolidated Framework for Implementing Research \((38\%)\) and RE-AIM \((39\%)\) models were the most commonly reported frameworks \((n = 16\) studies each). Promoting Action on Research Implementation in Health Services \((40, 41\)\) and the Theoretical Domains Framework \((42\)\) were each reported by 12 studies. Additional models that were used by multiple studies included diffusion of innovations \((43\) \((n = 8\)\) and the exploration, preparation, implementation, and sustainment model (EPIS, \(n = 5\) \((8\)\). Seven models were each reportedly used in two or three studies: Grol and Wensing’s implementation of change model \((44\)\); UK MRC Complex Interventions Framework \((45\)\); Normalization Process Theory \((46\)\); Chronic Care Model \((47\)\); Dynamic Sustainability Framework \((1\)\); Greenhalgh’s Model of Diffusion of Innovation in Health Organizations \((48\)\); and the Ottawa Model of Research Use \((49\)\). The remaining three models appearing only once in the sample.

### Additional Study Characteristics: Data Type, Study Location, and Funding Sources

One hundred twenty-nine studies \((61\%)\) used some combination of quantitative and qualitative data collection methods, and (since we excluded qualitative only studies) the remaining 39% \((n = 83\) utilized only quantitative methods. The majority of studies were conducted in the US \((n = 69, 33\%)\) or Canada \((n = 45, 21\%)\). There were 21 \((13\%)\) studies from Australia and 24 studies \((11\%)\) from the Netherlands. The remaining studies took place across Europe, Africa, and Asia. When considering funding sources, 183 \((86\%)\) of studies relied on regional or national agency contributions. Twenty-eight \((13\%)\) studies were funded by a foundation or internal funding, and 18 \((8\%)\) studies were funded by a regional, national, or agency, and four \((2\%)\) were funded by industry.
Several studies were funded by multiple types of funding, and as such, one study may be represented in more than one of these categories.

**DISCUSSION**

The current review found that of the included D&I studies from the protocol papers published in the journal *Implementation Science*, most are using cluster randomized trials or RCTs, although the use of RCTs has decreased. Though a number of other designs have been proposed to conduct D&I research (4, 50), these alternative designs may be under-represented in the current findings, and RCTs still predominate D&I literature (17). This is particularly noteworthy given the review included only protocol papers from the journal *Implementation Science*, which is likely more “open” to new/other types of D&I designs than other scientific journals. D&I studies are also being published in other journals, which may have an even lower rate of alternate design types. However, this field is still relatively new, and it may take time to see a more balanced distribution of study designs appear within peer-reviewed literature.

The increase in the variety of study designs used over time indicates that researchers are using alternative designs more frequently to answer different D&I research questions. As described by Aarons and colleagues, these questions take place across different phases of D&I research that include exploration, adoption/preparation, implementation to identify effective D&I strategies for improving program fidelity, and sustainment to examine strategies that promote maintained delivery or use of an EBI (8). Some designs may be more suited to answer particular research questions within each phase. For example, a comparative case study design is appropriate to identify a potentially effective implementation strategy to test in future research (51), while a cluster randomized stepped wedge design may be more appropriate when testing the effectiveness and sustainability of an implementation strategy (52). We could not code for this within our sample, as it is not always specified which phase researchers consider their research questions, but it is possible this is a factor in deciding which design to use.

Given the benefits of using a theory or framework to guide D&I research (53–57), it is surprising that the current review identified only 111 (52%) studies that described such grounding. Other reviews have also found low prevalence of theory and framework use (58–60), even though resources exist to help D&I researchers search for and identify appropriate theories or frameworks to guide their studies (61, 62). These studies may have a theoretical underpinning that was not articulated in the protocol. However, there is a need for wider use and reporting of theory and frameworks used, as they are known to increase the effectiveness of an implementation strategy (63), to understand the mechanism by which a program acts, and to promote replicability of studies.

Despite the significant benefits randomized trials can provide (i.e., internal validity), it is possible that their use may reduce external validity (64). Less traditional methods (e.g., multiple baseline design, phased implementation), which appear to be underutilized, provide enhanced flexibility and capacity to incorporate local context; these types of designs may additionally present more feasible options. Additionally, methods such as systems science and network analysis were not identified in the current review, but are growing in popularity in D&I research (65). However, it is possible that our inclusion criteria, particularly the requirement of a comparison group, may have excluded such methods.

While there has been an increase in the use of alternative designs, many researchers continue to rely on more traditional designs, such as RCTs, similar to a prior review of implementation studies specific to child welfare and mental health (30). There are likely many reasons researchers continue to utilize RCTs, including those designing and evaluating studies may perceive these as the best way to minimize selection bias. It is possible that our findings represent a dissemination issue, in that the use of alternate designs is gaining speed, but has been slow to spread through this newly developing field. To facilitate the spread of different and perhaps more appropriate designs and to assist investigators developing D&I studies, we have provided a guide for researchers making decisions about their study designs (Figure 3). This decision process begins with defining a research question (53–55), which determines whether the data needed should be quantitative, qualitative, or mixed. Once the research question and type of data are determined, it is important to consider whether it is possible and ethical to assign exposure and if the exposure can be assigned by group or by time. In the current review, the majority of studies reviewed included assigning exposures (n = 186, 88%). If assigning exposure randomly is ethical and practical, the study can be either experimental or if not, quasi-experimental; in the current review, 164 (77%) and 35 (17%) of included studies were randomized and quasi-experimental, respectively.

If randomization is not possible, then there are alternate ways to enhance the rigor of a design. For example, group equivalence at pre-test can be achieved by design factors such as matching or using matched controls (66). Other options to strengthen internal validity include multiple pre- and post-tests and/or removed then repeated interventions (9, 17, 50). In these types of studies, units can be randomized to different time periods (rather than only to groups), such as with stepped wedge designs. This helps account for time-related (e.g., history) threats to internal validity, etc., reducing threats to both internal and external validity (17, 23, 24). When assignment of exposure is possible, it is also important to consider the level at which exposure can/will be assigned (e.g., individual, organizational) and to address any clustering effect this might create through design, measurement, and analysis. Specific alternative designs do not appear in the figure; instead, opportunities for alternative designs exist within each category (e.g., randomize by time vs. condition).

Another alternative design when exposure is not ethical or possible is the observational design (67). The current review identified few studies using observational designs (n = 13, 6%). It is possible that our inclusion criteria may have led to this under-representation of observational designs, particularly cross sectional. Observational designs can vary considerably depending on whether data can be collected over time (i.e., longitudinal).
or at only one time (i.e., cross sectional). It might be possible to enhance the evaluative power of an observational study if data collection can be timed around implementation of an intervention to create a natural experiment. Observational designs might also be useful in pre-intervention phases, identifying prevalence rates, potential intervention points, hypothesized causal pathways, potential mediators, and acceptable implementation strategies (9, 67). The rigor of these studies can be enhanced with data collection at more time points, and the internal validity can be improved if measures with more reliability and validity evidence are used.

There are issues that cut across all of these decisions about study designs. It is beyond the scope of this paper to discuss all the potential decisions that might arise in study design, but three are of particular importance: context, study level, and use of a theory or framework. Context is the setting in which practice takes place and is particularly important in D&I research (68). Whether study sites are selected to represent a range of different organizations with respect to cultures, climates, readiness, or just selecting the sites that are most “ready” or amenable to the implementation effort is an important decision point with implications for interpretation of findings. Regardless of the decision around the study design, it is important that consideration of context be explicitly incorporated into the study, such as in site selection, as it can have important implications on whether an intervention is implemented properly and therefore can have its intended effects. Determining the level for assignment, intervention, and measurement, all have important implications (e.g., in the school setting: individual students, classrooms, schools, school districts). Within the coding scheme used for this review, it was sometimes difficult to identify these characteristics of studies, possibly because of differences across substantive areas. With the low use of theory in the studies for this review, there is an opportunity to strengthen future research with the use of theory that guides implementation and measurement and is articulated. Better reporting of study characteristics can promote replicability and translation of knowledge across disciplines.

Analytical methods may be utilized to account for these decisions (e.g., the use of multilevel modeling). Where possible researchers should be consistent in the levels at which they assign, intervene, and measure effects. Though this does not prevent bias, which can still exist even with consistency, it lessens the chance. These decisions also have important implications for sample size and statistical power (i.e., unlike in a clinical trial, where the sample may be at the level of the individual, D&I studies often require that units be the cluster organization, hospital, school, agency level) as well as analysis; when clustering is present, appropriate statistical measures must be employed.

Several issues in D&I research should influence the design choice. For example, if the intervention evidence is sound, it may not be necessary to re-establish effectiveness; rather, one may be more interested in tracking the fidelity of implementation. This often implies the need for knowledge about organizational
factors, including culture, climate, and readiness. In addition, measurement is important to consider. Whether or not measures exist to assess the factors in question, including the psychometric and pragmatic properties of these measures (69, 70), will inform design decisions (71, 72). The choice if a D&I design involves a series of trade-offs including some that are not addressed here, and these often balance scientific rigor with real-world circumstances (10). Specific examples of study designs proposed within this sample are available in Data Sheet S1 in Supplementary Material. Also, several examples have been presented in Data Sheet S2 in Supplementary Material based on the decision tree that detail some of these considerations, and Data Sheet S3 in Supplementary Material presents a compilation of resources available to support design choice.

This study has limitations worth noting. The first is that only protocol papers from one journal were included, and our sample may not be generalizable to all D&I research published in other journals or outside of a study protocol format. However, Implementation Science is on the forefront of the emerging field and likely represents a broad spectrum of studies being conducted in D&I research. Additionally, purely qualitative studies were not included in this review, and we did not code how qualitative and quantitative data were used within a study. Though few studies were excluded for this reason alone (n = 12), studies of this nature may demonstrate use of alternate study designs. Future research on the use of mixed methods within D&I work is needed to understand how types of mixed methods approaches are applied in D&I research (73). Another limitation of our sampling is our focus on research that is testing D&I strategies, thus leaving out a whole set of D&I studies that focus primarily on understanding the context including influences on professional and organizational behavior; these studies are often shorter in duration and likely from smaller grants, where investigators may not publish protocol papers. Further, our sample may have suffered from selection bias, as trials are most likely to be funded and to benefit from publishing a protocol paper. Thus, it might be expected that RCTs and cluster RCTs were common. We were also limited in coding what was presented in the protocol paper, and in some cases, during implementation of a study, some changes may be made that are not reported in the original protocol (e.g., addition of constructs from a different theory). Last, we did not code how the qualitative data were used within studies using both qualitative and quantitative data, i.e., parallel sequential or converted approaches (33).

In the face of national and international calls for accelerating the spread of EBIs, policies, and treatments, maximizing the utility of the results for D&I studies is essential. This includes findings with robust internal validity while maximizing external validity and those that are relevant to the variety of stakeholders involved in D&I research. Fortunately, the field has a suite of designs, including many alternatives to RCTs, which can help answer these calls.

CONCLUSION

While alternatives to the RCT (e.g., stepped wedge, adaptive designs) were employed in several studies, our review suggests that funded D&I research has largely mirrored clinical effectiveness research by primarily relying upon cluster RCTs and RCTs. However, alternative designs that offer researchers flexibility based on the context of their research and can maximize external validity are becoming more common. While the use of design approaches using qualitative and quantitative data sources appears to be prevalent in D&I research, there is a need for more use and reporting of D&I theory to guide future studies.

AUTHOR CONTRIBUTIONS

All authors made substantial contributions to conception and design, acquisition of data, and interpretation of data; SM and RT were involved in drafting the manuscript; MP, AR, AB, EK, EL, MP, BP, and RB have been involved in revising the manuscript critically for important intellectual content. All authors have given final approval of the version to be submitted and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

The authors would like to thank the members of the Washington University Network for Dissemination and Implementation Research (WUNDIR), and particularly Drs. Enola Proctor and Rebecca Lobb, for their help and guidance throughout the conception, interpretation, and presentation of the study. They would also like to thank Dr. David Chambers and Ms. Alexandra Morshed. Support for this project came from National Cancer Institute at the National Institutes of Health Mentored Training for Dissemination and Implementation Research in Cancer Program (MT-DIRC) (5R25CA171994-02) and the National Institute of Mental Health (5R25MH080916). Additional support came from the National Institute of Mental Health (5P30 MH068579, 5R25MH080916); the National Cancer Institute at the National Institutes of Health (5R01CA160327); the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK Grant Number 1P30DK092950); the National Institute on Drug Abuse of the National Institutes of Health (K12DA041449); the National Heart, Lung, and Blood Institute at the National Institutes of Health (3U01HL1339940251); the National Human Genome Research Institute at the National Institutes of Health (1R01HG00935101A1); Washington University Institute of Clinical and Translational Sciences grant UL1 TR000448 and KL2 TR000450 from the National Center for Advancing Translational Sciences; and grant funding from the Foundation for Barnes-Jewish Hospital. The content is solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/articles/10.3389/fpubh.2018.00032/full#supplementary-material.
REFERENCES


doi:10.1136/qhc.7.3.149


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Mazzucca, Tabak, Pilar, Ramsey, Bauman, Kryzer, Lewis, Padek, and Brownson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
A Protocol for a Feasibility and Acceptability Study of a Participatory, Multi-Level, Dynamic Intervention in Urban Outreach Centers to Improve the Oral Health of Low-Income Chinese Americans

Mary E. Northridge1*, Sara S. Metcalf2, Stella Yi3, Qiuyi Zhang2, Xiaoxi Gu1, Chau Trinh-Shevrin3 for the Implementing a Participatory, Multi-Level Intervention to Improve Asian American Health Research Team

1Department of Epidemiology and Health Promotion, College of Dentistry, New York University, New York, NY, United States, 2Department of Geography, The State University of New York at Buffalo, Buffalo, NY, United States, 3Department of Population Health, School of Medicine, New York University, New York, NY, United States

Introduction: While the US health care system has the capability to provide amazing treatment of a wide array of conditions, this care is not uniformly available to all population groups. Oral health care is one of the dimensions of the US health care delivery system in which striking disparities exist. More than half of the population does not visit a dentist each year. Improving access to oral health care is a critical and necessary first step to improving oral health outcomes and reducing disparities. Fluoride has contributed profoundly to the improved dental health of populations worldwide and is needed regularly throughout the life course to protect teeth against dental caries. To ensure additional gains in oral health, fluoride toothpaste should be used routinely at all ages. Evidence-based guidelines for annual dental visits and brushing teeth with fluoride toothpaste form the basis of this implementation science project that is intended to bridge the care gap for underserved Asian American populations by improving access to quality oral health care and enhancing effective oral health promotion strategies. The ultimate goal of this study is to provide information for the design and implementation of a randomized controlled trial of a participatory, multi-level, partnered (i.e., with community stakeholders) intervention to improve the oral and general health of low-income Chinese American adults.

Methods: This study will evaluate the feasibility and acceptability of implementing a partnered intervention using remote data entry into an electronic health record (EHR) to improve access to oral health care and promote oral health. The research staff will survey a sample of Chinese American patients (planned n = 90) screened at three outreach centers about their satisfaction with the partnered intervention. Providers (dentists and community health workers), research staff, administrators, site directors, and community advisory board members will participate in structured interviews about the partnered
intervention. The remote EHR evaluation will include group adaptation sessions and workflow analyses via multiple recorded sessions with research staff, administrators, outreach site directors, and providers. The study will also model knowledge held by non-patient participants to evaluate and enhance the partnered intervention for use in future implementations.

Keywords: implementation science, feasibility, acceptability, oral health, dental care, health equity, urban health, Chinese American

INTRODUCTION: BACKGROUND AND RATIONALE

In 2010, oral health conditions affected 3.9 billion of 6.9 billion people worldwide, or over half (57%) of the global population (1). Indeed, untreated tooth decay (dental caries) was the single most prevalent and severe periodontitis (periodontal disease) was the sixth most prevalent of 291 oral and general health conditions studied (1). The burden of unmet oral health needs on quality of life is substantial, especially for populations with fewer resources. Globally, disability adjusted life years (DALYs) lost due to dental caries and periodontal disease is considerable, especially in China and India (Figure 1), but also for disadvantaged populations worldwide.

The reasons for the documented oral health and health care disparities between countries and populations are complicated (2). They are now being understood in terms of ecological models (3, 4) which posit that factors at multiple levels—in the case of our project, community (level 4), site and provider (level 3), family (level 2), and patient (level 1)—influence disparities in access to and quality of services (Figure 2).

As per our multi-level approach, interventions that address factors at multiple levels may be more effective than those that target a single level (5). For instance, remote electronic health record (EHR) data entry at the institutional level will permit tracking and evaluation of our multi-level intervention at the community, site, provider, and patient levels, in accordance with the implementation strategy to change record systems. This will better enable us to determine whether or not our multi-level intervention results in, e.g., improved patient care at the individual level and enhanced oral health at the community level.

In the United States, oral diseases ranging from dental caries to oral cancers cause pain and disability for millions of US children and adults (6). This “silent epidemic” has disproportionately severe impacts on marginalized populations, including people from historically disadvantaged backgrounds and racial/ethnic minorities (7). Further, a thorough oral examination can detect signs of general health problems, including nutritional deficiencies, immune disorders, injuries, and certain cancers, with referrals to apt health care providers where indicated (7). Relatedly, and of critical, but underappreciated importance, there is a segment of the population that visits a dental provider, but not a primary medical provider each year. Using data collected as part of the Medical Expenditure Panel Survey (MEPS), Strauss et al. found that of the 24.1% of adults who do not access general outpatient care, 23.1% visited a dentist (8). Hence, chairside screening for chronic conditions, such as diabetes and hypertension with referrals to primary care providers where indicated has the potential to identify persons who are unaware of their disease status, improve patient health, and lower health care costs (9, 10).

As a result, both oral and general health conditions may be prevented by regular dental visits (6). Accordingly, Healthy People 2020 Leading Health Indicator (meaning that it is a high priority, evidence-based health issue) is Oral Health-7 (OH-7), namely: increase the proportion of children, adolescents, and adults who used the oral health care system in the past years (6). Unfortunately, the US population as a whole has been moving away from the target of 49.0%, from a baseline of 44.5% in 2007 to 42.1% in 2012. For the Asian-only subgroup, the same trend exists, but at a lower prevalence over time, from a baseline of 41.3% in 2007 to 38.2% in 2012 (6).

The oral health benefits of fluoride have been well known for more than 70 years (11). Specifically, fluoride reduces the risk of dental caries in both children and adults through a variety of mechanisms, including: incorporating into enamel before teeth erupt; preventing demineralization and enhancing remineralization of teeth; and inhibiting bacterial activity in dental plaque (7, 12, 13). The Centers for Disease Control and Prevention recommends drinking fluoridated water if it is available and using fluoride toothpaste (12). The American Dental Association (ADA) recommends brushing teeth for 2 min twice a day with a soft-bristled toothbrush and fluoride toothpaste as part of a complete dental care routine (14). A video on proper brushing technique is also available on the ADA website (14).

Rationale for Prioritizing Asian American Oral Health Research

There are at least four compelling reasons to prioritize Asian American oral health research. First, Asian Americans are the fastest growing minority group in the United States, increasing in size by 43.3% between 2000 and 2010, more than four times faster than the total US population. Nationwide, there are nearly 14.6 million Asian Americans, representing approximately 4.8% of the US population, of whom more than 60% are foreign born and more than 30% have limited English proficiency (15). Second, studies have shown that current research and policy practices give rise to erroneous conclusions about Asian American health due to omission from data collection efforts, aggregation across Asian subgroups, and extrapolation of results from one Asian subgroup to another (16, 17). The preponderance of the limited research is based on West Coast populations, leading to a dearth of understanding of substantial and emerging Asian American communities in the Midwest.
FIGURE 1 | The burden of oral conditions worldwide as measured by disability adjusted life years lost due to tooth decay and periodontal disease, 2010. Source: Benzian and Williams (1). Printed with permission.

FIGURE 2 | This graphic is derived from the conceptual model, Factors that influence disparities in access to care and quality of health care services, by level created from the analysis of findings from systematic reviews published as: Purnell et al. (4).
Northeast, and Southwest. Third, the “model minority” stereotype of Asian Americans as “wealthier, wiser, and healthier” than other racial/ethnic minority populations undermines the significance of health disparities experienced among and within Asian American communities and the need to devote resources to mitigate those disparities (18). Fourth, findings of community health resources and needs assessments that sampled underserved and hard to reach New York, NY Asian immigrant populations consistently found that oral disease was a top concern for Chinese and Sikh South Indian populations (19–21). Fully 22% of Chinese respondents ranked oral or dental health as their top health concern, 68% rated their oral health as poor or fair, and only 53% reported having received an oral/dental health check-up in the past year (20).

In separate and related research conducted in New York, NY, the self-reported frequency of visiting a dentist in the past year for Chinese respondents ranged dramatically depending upon the sample studied, from 15.7% among participants recruited from recent immigrant enclaves (22) to 89.3% among participants recruited from throughout the five boroughs who were low-income, but highly educated (23). In a prospective study that examined racial/ethnic differences in periodontal disease among participants recruited from six US sites, Chinese participants displayed the highest prevalence of self-reported periodontal disease (39.8%), followed by Blacks (32.0%) and Whites (26.0%), with Hispanics displaying the lowest prevalence (17.4%) (24). Finally, a study conducted in Washington, DC among Chinese, Korean, and Vietnamese Americans found that less acculturated Asian Americans were less likely to receive physical, dental, and eye examinations than those who were more acculturated (25). This may be because there is less emphasis on preventive health care in Asian cultures, or because barriers to health care access may amplify the reluctance for preventive care among Asian Americans (25).

Adaptation of the Sikh American Families Oral Health Promotion Program

Using a community-based participatory research (CBPR) approach, UNITED SIKHS, a community-based organization that pursues projects for the spiritual, social, and economic empowerment of underprivileged and minority communities (26), New York University School of Medicine (NYU Medicine), City University of New York Prevention Research Center (27), and NYU College of Dentistry (NYU Dentistry) (28) developed, implemented, evaluated, and disseminated the Sikh American Families Oral Health Promotion Program (29). Several of the implementation strategies identified by Powell et al. were found to be effective in this earlier effort will be adapted for the proposed project, including conduct a local needs assessment, involve patients and family members, use to train-the-trainer strategies to provide hands-on instruction on proper brushing techniques to community educators, and use a community advisory board (CAB) to provide input and advice on implementation efforts (30). Findings were for Sikh participants with no dental insurance prior to program enrollment (n = 58), 81.0% credited the program with helping them obtain insurance for them or their children; for participants with no dentist prior to program enrollment (n = 68), 92.6% credited the program with helping them or their children find a local dentist (29).

Systems Modeling to Understand Dynamic Complexity and Simulate Alternate Scenarios

The term “systems science” is used to refer to the big picture of problem solving, where the problem space is conceptualized as a system of interrelated component parts (31). Both the coherent whole of the system and the relationships among the component parts are critical to the system, as they give rise to emergence, meaning much coming from little (32). Note that emergence occurs when even a relatively simple system generates unexpected amounts of complexity, which cannot be understood without the ability to simulate (32).

In order to improve our mental models of the real world, system scientists have developed and leveraged methods, such as system dynamics (SD), agent-based modeling (ABM), geographic information science, and social network simulation. The practice of systems science modeling is situated amidst an ongoing process of observing the real world, formulating mental models of how it works, setting decision rules to guide behavior, and from these heuristics, making decisions that in turn affect the state of the real world (33).

Interventions often fail or even worsen the problems they are intended to solve due to a lack of understanding of real world structures and dynamic complexity. Among the benefits of systems modeling are iterative practice, participatory potential, and possibility thinking. Best principles and recommendations for advancing implementation science through systems science modeling are summarized below (Table 1), based upon the seminal contributions of thought leaders in the field (34).

As part of a body of research to understand the complex set of causal pathways and time delays that compound health

<table>
<thead>
<tr>
<th>Best principles</th>
<th>Recommendations</th>
<th>Key references</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Model the problem, not the system</td>
<td>Conduct formative research; construct models collaboratively in interdisciplinary teams</td>
<td>Serman (53)</td>
</tr>
<tr>
<td>2. Pay attention to what is important, not just what is quantifiable</td>
<td>Use qualitative data to derive causal relationships; be guided by deep thinking and multiple perspectives</td>
<td>Meadows (35)</td>
</tr>
<tr>
<td>3. Leverage the utility of models as boundary objects</td>
<td>Create modifiable and accessible representations of models; build trust by representing local knowledge</td>
<td>Black (36)</td>
</tr>
<tr>
<td>4. Adopt a portfolio approach to model building</td>
<td>Work in parallel to develop separate, but related models in diverse ways; encourage exploration</td>
<td>Metcalf et al. (37)</td>
</tr>
</tbody>
</table>
inequities over the life course, our research team developed, refined, and tested a portfolio of systems science models that originated in the ElderSmile program in northern Manhattan (38). Despite the time and resources required to ensure a participatory approach to group model building that elicits the knowledge of all team members across disciplines and fields, the simulation models devised may more accurately reflect real world conditions and possibilities. If so, in the end, time and resources will have been well spent in the service of running virtual experiments that may more effectively direct program enhancements and policy changes that improve the health and well-being of disadvantaged adults, and may be adapted for other populations and locales.

Statement of Compliance
The feasibility and acceptability study will be conducted in accordance with the International Council on Harmonization guidelines for Good Clinical Practice (ICH E6), the Code of Federal Regulations on the Protection of Human Subjects (45 CFR Part 46), and the National Institute of Dental and Craniofacial Research (NIDCR) Clinical Terms of Award. All personnel involved in the conduct of this study have completed human subject’s protection training.

Potential Risks and Benefits
Potential Risks
Despite the enhanced computer security, patient data may be at a higher risk for computer hacking due to remote data entry, leading to a loss of medical record confidentiality.

Potential Benefits
Outreach center clients/patients may benefit from the intervention, including the translated and culturally customized literature, by being influenced to pursue dental care and conduct twice daily dental hygiene using evidence-based products and procedures.

Objectives
The ultimate goal of this study is to provide information for the design and implementation of a randomized controlled trial of a participatory, multi-level, partnered (i.e., with community stakeholders) intervention to improve the oral and general health of low-income Chinese American adults. Toward this end, this study has three objectives.

Primary
- To evaluate and enhance the feasibility and acceptability of a partnered intervention designed to improve oral health for low-income, urban Chinese American adults at three community sites.

Secondary
- To evaluate and enhance the feasibility and acceptability of using remote entry features of EHR software at NYU Dentistry to enter patient information at three Chinese American community sites.

- To model knowledge held a priori by non-patient participants about factors that influence access to oral health care and care-seeking behaviors among low-income, urban Chinese American adults, in order to enhance the intervention during and/or after the study for use in future implementations

METHODS: FRAMEWORKS, SETTING, POPULATION, AND SPECIAL TERMS

Theory-Driven Implementation Frameworks
Our study design to implement remote EHR data entry and tracking and a partnering package of evidence-based intervention strategies in diverse Chinese American community outreach sites is guided by two complementary, multi-level frameworks: Consolidated Framework for Implementation Research (CFIR) (39, 40) and Implementation Outcomes Framework (IOF) (41, 42). Specifically, CFIR provides a menu of constructs that have been associated with effective implementation and have been used in a range of applications, including our own oral health research (29, 43–46). For the proposed project, the five domains and the associated constructs that we are particularly interested in exploring are: (1) intervention = partnered EHR enhanced community outreach (adaptability, cost); (2) inner setting = NYU Dentistry Local Community Outreach Programs and clinics (implementation climate, relative priority); (3) outer setting = Chinese American outreach sites (patient needs and resources); (4) characteristics of individuals involved = champion (Dr. Wolff), implementation leaders (Drs. Schenkel and Perelman), external change agents (CAB members and site directors), researchers, dental providers, family members, and patients (self-efficacy); and (5) process of implementation (planning, engaging, reflecting, and evaluating). A final critical component of CFIR is the process of adaptation of the intervention for diverse partnering sites.

Implementation Outcomes Framework is clear in distinguishing implementation outcomes (acceptability, adoption, implementation cost, sustainability), service outcomes (effectiveness, equity), and client outcomes (satisfaction), all of which we intend to assess. While IOF provides an evaluation outcomes framework that organizes the multiple facets that affect implementation of new interventions, CFIR provides a framework for understanding the multiple domains that influence implementation and adoption of these interventions.

Finally, because our proposed intervention is both multi-level and dynamic with numerous involved constructs, we intend to model the knowledge gained about factors at the community, site, provider, family, and patient levels to improve oral health using a participatory group modeling approach. We will leverage the power and flexibility of software programs, such as AnyLogic1 and Vensim,2 to construct simulation models that enable integration of different structural components of models: agents, social networks, geographic information system data, and stock-flow SD.

1https://www.anylogic.com/.
2http://vensim.com/.
Rationale for the Selection of the Setting and Population

The setting for this project is the NYU Dentistry Local Community Outreach Programs (20). NYU Dentistry conducts local community outreach in two ways. First, all dental students must complete 4-month long community-based rotations, in which they spend 1-day per week providing direct patient care under faculty supervision in 1 out of 7 locations in four boroughs of New York, NY. A second type of outreach, which is the focus of this research, takes place in dozens of locations throughout New York, NY; it is an entirely voluntary effort shared by faculty preceptors and students. In 2015, the number of volunteer community events reached 126. The volunteer screening events are held an average of three times per week, on weekdays and weekends, with 6–8 students typically taking part in each event. Although they do not directly treat patients at these sites, students refer many of them to NYU Dentistry. To encourage patients to visit a dentist, students provide each patient screened with a voucher worth $205.00 for oral health care at NYU Dentistry to cover his/her comprehensive oral examination, treatment plan, and prophylaxis at no charge and with no co-payment required.

In discussions with Dr. Schenkel, the program leader, we learned that turnout at Chinese sites was especially high, affirming the findings of a recent Chinese Community Health Resources and Needs Assessment conducted by NYU Medicine, where oral health was identified as a top concern (20). Dr. Schenkel also expressed a need for data collection and analysis to evaluate the program. With guidance from Dr. Wolff, the project champion, Dr. Perelman, the IT leader, was recruited to plan, implement, and evaluate remote EHR data entry and tracking at community sites.

Populations/Units of Analysis for the Feasibility and Acceptability Study

As we are utilizing a CBPR approach, several different populations/units of analysis are included in this feasibility and acceptability study:

- In accordance with CBPR principles, we will establish a CAB comprised of eight members to guide all aspects of this study. We are currently recruiting CAB members from the following local partners: Chinatown YMCA, Chinatown Planning Council Manhattan, Hamilton-Madison House, and the NYU Medicine Clinical and Translational Science Institute Integrating Special Populations Program.
- The research staff (n = 10) comprises of three Multiple Principal Investigators (MPIs) (Drs. Northridge, Trinh-Shevrin, and Metcalfe), the three Co-Investigators (Co-Is) at NYU Medicine (Drs. Troxel, Islam, and Yi), two Project Coordinators (TBN), and two modelers (Ms. Zhang and TBN).
- The NYU administrators (n = 3) include both Co-Is at NYU Dentistry (Drs. Schenkel and Wolff) and an IT Specialist (Dr. Perelman).
- The providers (n = 8) include faculty dentists at NYU Dentistry who participate in local community outreach events (n = 6) and bilingual (English and Mandarin Chinese) community health workers (CHWs) (n = 2).
- Outreach site directors at each of three participating community outreach centers (n = 3) serving low-income Chinese American populations in New York, NY.
- EHR patient participants. At least 50 low-income Chinese American adult patients will undergo a dental screening at each of three community outreach centers (n = 150), and have their data remotely entered into the NYU Dentistry EHR.
- Interview patient participants. Approximately 30 Chinese American patients screened at each of three community outreach centers will be enrolled in the study as participants (n = 90) to complete interviews about their satisfaction with the partnered intervention and their use of dental services and evidence-based oral health behaviors.

Special Terms

Terms with a special meaning regarding this protocol are explained next.

- The partnered intervention consists of the following four evidence-based strategies: (1) written agreements of collaboration for dental screening at the site level; (2) culturally tailored and language-specific adaptation of materials at the community and site levels; (3) demonstrations with role-playing of proper brushing with fluoride toothpaste and flossing techniques at the site and provider levels; and (4) CHW follow-up with patients about oral health care receipt and dental hygiene behaviors at the family and patient levels.
- Patient participants refer to study participants who are also patients screened at the urban outreach centers. Patient participants are first enrolled as EHR patient participants. A subset of the EHR patient participants also complete exit interviews after the intervention at urban outreach centers and are referred to as interview patient participants (Figure 3).
- Non-patient participants refer to all other study participants who are not patients screened at the urban outreach centers, including CAB members, research staff, NYU administrators, providers (dentists and CHWs), and outreach site directors (Figure 4).

METHODS: OUTCOMES, DESIGN, ENROLLMENT, AND WITHDRAWAL

Study Outcome Measures

The primary and secondary outcome measures for this feasibility and acceptability study are provided below.

Primary

Patient satisfaction with the partnered intervention components based on exit interviews.

Secondary

Table 2 provides the series of secondary outcome measures associated with the objectives of this feasibility and acceptability study.
Other important outcomes of this feasibility and acceptability study include:

- Work flow analysis of the interviews of research staff, NYU administrators, and providers (dentists and CHWs) is a secondary measure designed to evaluate and refine the use of the remote EHR.
- Knowledge, expertise, insights, and opinions of the research staff, NYU administrators, providers (dentists and CHWs), outreach site directors, and CAB members about factors that influence access to oral health care and care-seeking behaviors among low-income, urban Chinese American adults are secondary measures designed to be used in simulations to represent dynamics at multiple levels (patient, family, provider, site, and community).
- The simulation modeling platform is a research product of this study that will be used to experiment with strategies to promote preventive and restorative care through regular dental visits and self-efficacy among Chinese Americans.

**Study Design**

This feasibility and acceptability study will be conducted at three community outreach centers serving an urban, low-income Chinese American population. The study will evaluate the feasibility and acceptability of implementing a partnered intervention to improve the oral and general health of low-income, urban Chinese American adults and of using remote entry into an EHR. The evaluation will include group adaptation sessions and workflow analyses of the EHR implementation, involving multiple recorded sessions with NYU administrators, providers (dentists and CHWs), outreach site directors, and research staff. The study will also model *a priori* knowledge held by non-patient participants to evaluate and enhance the intervention during and/or after the study for use in future implementations (Figure 5).

Approximately 50 patient participants who self-identify as Chinese American from each of three outreach centers (*n* = 150) will be consented to allow the entry of their data (e.g., demographic information, medical history, receipt of oral health care visits, dental hygiene behaviors, and health and health care measures) into the remote EHR by authorized NYU Dentistry staff (EHR patient participants). Of these 150 EHR patient participants,
TABLE 2 | List of secondary outcome measures, with their corresponding constructs, levels of analysis, and data sources.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Levels of analysis</th>
<th>Measures (Quantitative/Qualitative)</th>
<th>Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation</td>
<td>Provider site</td>
<td>Acceptability Provider institution, uptake and utilization of remote EHR and partnering components</td>
<td>Exit interviews with patients; semi-structured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extent to which the remote EHR entry and partnered intervention model are compatible with resources and training</td>
<td>interviews</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Costs Institution Intervention and implementation costs, including investment, supply, and</td>
<td>Observation; semi-structured interviews; EHR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>opportunity costs</td>
<td>Semi-structured interviews; EHR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feasibility Provider site, extent to which the remote EHR entry and partnered intervention model are</td>
<td>Semi-structured interviews; EHR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fidelity Provider Site, provider adherence to program protocol and quality of delivery</td>
<td>Community health worker logs; self-report</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sustainability Institution Site, sustained remote EHR use at outreach events and partnering package of</td>
<td>Semi-structured interviews; EHR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interventions</td>
<td></td>
</tr>
<tr>
<td>Service outcome</td>
<td>Community provider</td>
<td>Equity Provider site, support from community partners, providers (including NYU Dentistry), family</td>
<td>Baseline survey; follow-up patient interviews;</td>
</tr>
<tr>
<td>characteristics</td>
<td>family patient</td>
<td>members, and patients to direct resources to less well-served and less well-studied populations</td>
<td>semi-structured interviews; EHR</td>
</tr>
</tbody>
</table>
|                  |                    | Study Enrollment Patient Participants
Outreach center patients will be enrolled in either or both of two groups.

**Group 1**
Approximately 50 patients from each of three centers ($n = 150$) will be consented to allow their data to be entered via the remote EHR. These EHR patient participants must meet all of the following criteria:

1. Greater than or equal to 21 years of age.
2. Self-identify as being of Chinese ethnicity.
3. Live in any of the five boroughs of New York, NY and visit a participating outreach center.
4. Able and willing to provide informed consent and participate in an exit interview and a follow-up interview.

**Group 2**
Approximately 30 patients from each of three centers ($n = 90$) will be consented to participate in an exit interview and a follow-up interview. These interview patient participants must meet all of the following criteria:

1. Greater than or equal to 21 years of age.
2. Self-identify as being of Chinese ethnicity.
3. Live in any of the five boroughs of New York, NY and visit a participating outreach center.
4. Able and willing to provide informed consent and participate in an exit interview and a follow-up interview.

**Non-Patient Participants**
Non-patient participants will be enrolled into either or both of two groups.

**Group 1**
Approximately 20 research staff, NYU administrators, outreach center directors, and providers (dentists and CHWs) will be enrolled to participate in interviews about the partnered intervention and/or remote EHR. These non-patient participants must meet all of the following criteria:

1. Greater than or equal to 18 years of age.
2. Be employed or volunteers at participating outreach centers or employed at NYU.
3. For CHW-staff, speak and read Mandarin Chinese.
4. Able and willing to provide informed consent.

**Group 2**
Approximately 32 non-patient participants research staff, NYU administrators, CAB members, outreach site directors, and providers (dentists and CHWs) will be enrolled to participate in interviews and a group model-building workshop to inform model development by sharing their knowledge about factors that influence access to oral health care and care-seeking behaviors among low-income, urban Chinese American adults.
among low-income, urban Chinese American adults. These individuals must meet all of the following criteria:

1. Greater than 18 years of age.
2. Be employed or volunteers at participating outreach centers or employed at NYU.
3. Able and willing to provide informed consent.

**Subject Exclusion Criteria**

Individuals meeting any of the following criteria will not be enrolled as either EHR patient participants or interview patient participants:

1. Have an acute or terminal illness or a serious mental illness or any other severe health condition(s) that might preclude visiting an oral health care provider.
2. Are currently participating in another oral health study.

Individuals meeting any of the following criteria will not be enrolled to complete the interviews about the partnered intervention or remote EHR or to provide input to the knowledge modeling activities:

1. Staff in functional areas that do not directly service patients (e.g., custodial staff).

A patient participant may participate in either the EHR patient participant group only or both patient participant groups (interview patient participants are a subset of EHR patient participants). A non-patient participant may participate in any or all of the non-patient participant data collection activities. Co-participation in activities by any subject is not required.
Strategies for Recruitment and Retention

Five Chinese American community-based organizations have already volunteered to participate in this study:
- Asian Americans for Equality.
- Chinatown YMCA, a Branch of the YMCA of Greater New York.
- Chinese American Planning Council.
- Coalition for Asian American Children and Families.
- Hamilton-Madison House.

The three outreach centers for this study will be selected from among the affiliated outreach centers of these organizations. Clients from the outreach centers will be recruited by research staff working with three Chinese American community sites in New York, NY.

Electronic health record patient participants (of whom certain individuals are also interview patient participants) will receive a voucher worth $205 for oral health care at NYU Dentistry to cover his/her comprehensive oral examination, treatment plan, and prophylaxis at no charge and with no co-payment required as compensation for participation.

Non-patient participants—research staff, NYU administrators, providers (dentists and CHWs), outreach site directors, and CAB members—will receive no monetary compensation for their participation in the study, over and above their salaries/stipends.

Study Withdrawal

Reasons for Withdrawal

Any of the various participants (i.e., CAB members, outreach site directors, EHR patient subjects, interview patient subjects, research staff, NYU administrators, dentists, and CHWs) may withdraw from the study at any time. Patients will have the right to refuse to participate without any compromise of their health or dental services. Also, if a participant is uncomfortable during an interview or survey administration, he/she may stop at any time without penalty.

Handling of Subject Withdrawals or Subject Discontinuation of Study Intervention

If an EHR patient participant withdraws consent, no further data from that patient will be entered into the EHR for that participant. Depending on the nature of the request to withdraw, it may be necessary to remove existing data for that patient from the EHR.

Premature Termination or Suspension of Study

This study has no explicit stopping rules. The study may be suspended or prematurely terminated if there is sufficient reasonable cause. Written notification, documenting the reason for study suspension or termination, will be provided by the suspending or terminating party to the MPIs and/or the NIDCR, as applicable. If the study is prematurely terminated or suspended, the MPIs will promptly inform the Institutional Review Boards (IRBs) and provide the reason(s) for the termination or suspension.

Circumstances that may warrant termination include, but are not limited to:
- Determination of an unexpected, significant, or unacceptable risk to participants.
- Insufficient adherence to protocol requirements.
- Data that are not sufficiently complete and/or evaluable.

METHODS: INTERVENTION, TRAINING, SCHEDULE, AND ACTIVITIES

Administration of Intervention

Initially, a CAB will be established to guide all aspects of the study.

Partnered Intervention

Our partnering package of interventions builds upon the evidence-based practices of the NYU Dentistry Local Community Outreach Programs and the results of our pilot study in the Sikh American community (20), and aligns with the implementation strategies from the Expert Recommendations for Implementing Change (EPIC) project (30). We will work closely with the directors of three Chinese American sites to create written agreements of collaboration that outline the roles and responsibilities of the investigative team and the sites. An integral part of effectively implementing oral health activities with low income, racial and ethnic minority, and immigrant populations is to develop program materials that are specific to the local community (47). Training lay individuals of the same cultural and linguistic background as participants, e.g., CHWs through train-the-trainer techniques, has been found to be an acceptable approach for delivering culturally appropriate, community-based oral health interventions, as well as for recruiting participants into interventions through community and social networks (48–51). CHWs have been found to be effective in providing dental education and counseling (52), leading interactive demonstrations of brushing with fluoride toothpaste and flossing (53, 54), and improving access to dental care through dental coverage and linkage to local dentists (20).

Development of Culturally Tailored and Language-Specific Materials

The CAB will be responsible for reviewing existing program materials as an integral part of adapting them for the local Chinese American population. This will entail a multi-step process. Existing English and simplified Mandarin Chinese language materials will be presented to the CAB. Dr. Yi will then lead a guided discussion structured around the 4 P’s of social marketing. For product, CAB members will be asked if the materials encourage prevention of oral conditions through regular dental visits and brushing with fluoride toothpaste. For price, CAB members will be asked how much it will cost a person to take on the desired dental care behavior and the results of our pilot study in the Sikh American community (20), and aligns with the implementation strategies from the Expert Recommendations for Implementing Change (EPIC) project (30). We will work closely with the directors of three Chinese American sites to create written agreements of collaboration that outline the roles and responsibilities of the investigative team and the sites. An integral part of effectively implementing oral health activities with low income, racial and ethnic minority, and immigrant populations is to develop program materials that are specific to the local community (47). Training lay individuals of the same cultural and linguistic background as participants, e.g., CHWs through train-the-trainer techniques, has been found to be an acceptable approach for delivering culturally appropriate, community-based oral health interventions, as well as for recruiting participants into interventions through community and social networks (48–51). CHWs have been found to be effective in providing dental education and counseling (52), leading interactive demonstrations of brushing with fluoride toothpaste and flossing (53, 54), and improving access to dental care through dental coverage and linkage to local dentists (20).

Development of Culturally Tailored and Language-Specific Materials

The CAB will be responsible for reviewing existing program materials as an integral part of adapting them for the local Chinese American population. This will entail a multi-step process. Existing English and simplified Mandarin Chinese language materials will be presented to the CAB. Dr. Yi will then lead a guided discussion structured around the 4 P’s of social marketing. For product, CAB members will be asked if the materials encourage prevention of oral conditions through regular dental visits and brushing with fluoride toothpaste. For price, CAB members will be asked how much it will cost a person to take on the desired dental care behavior and the results of our pilot study in the Sikh American community (20), and aligns with the implementation strategies from the Expert Recommendations for Implementing Change (EPIC) project (30). We will work closely with the directors of three Chinese American sites to create written agreements of collaboration that outline the roles and responsibilities of the investigative team and the sites. An integral part of effectively implementing oral health activities with low income, racial and ethnic minority, and immigrant populations is to develop program materials that are specific to the local community (47). Training lay individuals of the same cultural and linguistic background as participants, e.g., CHWs through train-the-trainer techniques, has been found to be an acceptable approach for delivering culturally appropriate, community-based oral health interventions, as well as for recruiting participants into interventions through community and social networks (48–51). CHWs have been found to be effective in providing dental education and counseling (52), leading interactive demonstrations of brushing with fluoride toothpaste and flossing (53, 54), and improving access to dental care through dental coverage and linkage to local dentists (20).

For development of culturally tailored and language-specific materials, the CAB will be responsible for reviewing existing program materials. This will entail a multi-step process. Existing English and simplified Mandarin Chinese language materials will be presented to the CAB. Dr. Yi will then lead a guided discussion structured around the 4 P’s of social marketing. For product, CAB members will be asked if the materials encourage prevention of oral conditions through regular dental visits and brushing with fluoride toothpaste. For price, CAB members will be asked how much it will cost a person to take on the desired dental care behavior and the results of our pilot study in the Sikh American community (20), and aligns with the implementation strategies from the Expert Recommendations for Implementing Change (EPIC) project (30). We will work closely with the directors of three Chinese American sites to create written agreements of collaboration that outline the roles and responsibilities of the investigative team and the sites. An integral part of effectively implementing oral health activities with low income, racial and ethnic minority, and immigrant populations is to develop program materials that are specific to the local community (47). Training lay individuals of the same cultural and linguistic background as participants, e.g., CHWs through train-the-trainer techniques, has been found to be an acceptable approach for delivering culturally appropriate, community-based oral health interventions, as well as for recruiting participants into interventions through community and social networks (48–51). CHWs have been found to be effective in providing dental education and counseling (52), leading interactive demonstrations of brushing with fluoride toothpaste and flossing (53, 54), and improving access to dental care through dental coverage and linkage to local dentists (20).

Frontiers in Public Health | www.frontiersin.org February 2018 Volume 6 Article 29
Procedures for Training Interventionists and Monitoring Intervention Fidelity

Both CHWs in this study were formerly trained as CHWs, have a history of engaging in health promotion with the Chinese American community, and are bilingual in English and Mandarin Chinese (the primary dialect of participating community outreach sites). Specifically, the project CHWs were previously trained in a core competency program that employed diverse training methods, guided by adult learning principles and popular education philosophy. We will further train the project CHWs in the oral health promotion demonstration protocol and on oral health services and programs available at local clinics and hospitals. The investigative team and CAB members will also collect, assess, and deliver to the CHWs updated information regarding health and dental insurance and access to oral health programs available for low-income and immigrant communities. NYU Dentistry investigators and staff will train the CHWs using models on evidence-based oral health practices, stressing the importance of drinking fluoridated water and brushing teeth for 2 min twice a day with a soft-bristled toothbrush and fluoride toothpaste (“painting the teeth with fluoride”). This train-the-trainer model will promote peer support and allow the project to be replicated and sustained across settings. Interactive educational techniques will be integrated into the demonstrations. As in our pilot work (20), hands-on instruction will be provided on proper brushing and flossing techniques, culturally tailored health promotion methods (i.e., preparing healthy traditional meals and using the plate method to determine the proper balance and size of portions), and goal-setting skills. Trainees will then demonstrate the presented procedures back to the trainers using models.

At the end of the training, the CHWs will collaborate in groups to practice delivering short excerpts from the curriculum to their peers and project team members, with the trainers providing comments and assistance. Approximately, 1-month before the CHW training is complete, the curriculum will be pilot tested to ensure its cultural appropriateness with patients. Two mock educational sessions and a final examination of knowledge and evaluation of trial encounters with mock participants will be conducted with project investigators and CAB members. Individuals who score below the threshold level of knowledge regarding oral health promotion will receive intensive 1-on-1 tutoring and be required to take a second examination of knowledge. Quality assurance controls will be built into the intervention. Drs. Northridge and Yi will meet with the CHWs on an approximately bi-weekly basis to ensure that the model components are being consistently applied. Each CHW will keep a log of activities and communication around their follow-up of patients. These logs will be reviewed as necessary to evaluate the type and nature of communications between the CHWs and their assigned study participants.

Study Schedule

The study will extend for approximately 1 year. An approximate timeline for implementation of the various aspects of this study is provided below in Table 3.

Study Activities by Phase

Pre-Intervention Activities

- Create written agreements of collaboration that outline the roles and responsibilities of the investigative team and the outreach sites.
- Develop culturally tailored and language-specific program materials.
- Train CHWs in the intervention protocol.
- Conduct a workflow analysis of remote EHR data entry at Chinese American outreach centers.

Outreach Center Activities

- Health Insurance Portability and Accountability Act (HIPAA) certified research staff members or volunteers will explain the consent form, confidentiality agreement, and liability release to each potential patient participant and obtain his/her signature.

<table>
<thead>
<tr>
<th>Study activities</th>
<th>Months 1–6</th>
<th>Months 7–12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruitment of Community Advisory Board members and outreach sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review, update, and finalize community health worker (CHW) training materials/ Gain Institutional Review Board (IRB) approvals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train CHWs, evaluate fidelity, retrain as necessary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implement partnered intervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exit and 1-month follow-up interviews of enrolled interview patients who participated in the partnered intervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi-structured interviews of non-patient participants regarding partnered intervention (including materials review prior to implementation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workflow analysis for remote electronic health record (EHR) entry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pilot testing and live-usability for remote EHR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi-structured interviews of non-patient participants in EHR implementation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi-structured interviews of non-patient participants to inform model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group model-building workshop</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of simulation modeling platform</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data collection activity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The gray shades denote the time period in which the activities will take place.
Remote EHR Data Collection and Dental Screening:

- Authorized NYU Dentistry personnel will directly enter the demographic information of patient participants into the customized EHR using the study-specific SCREENING option.

EHR patient participants will complete a short paper intake form that will be scanned directly into a designated SCREENING folder in the EHR regarding medical conditions that could affect, or be affected by, their oral health (Yes/No checklist) and a brief questionnaire on self-reported receipt of a dental visit in the past year and dental hygiene behaviors (Oral Health Survey). These questions will be based on the World Health Organization Oral Health Questionnaire for Adults: (1) How often do you clean your teeth? (Never, Once a month, 2–3 times a month, Once a week, 2–6 times a week, Once a day, Twice or more a day); (2) Do you use any of the following to clean your teeth? Toothbrush, Wooden toothpicks, Plastic toothpicks, Thread (dental floss), Charcoal, Chewstick/miswak, and Others (Please specify); (3) For those who nominate Toothbrush: what type of toothbrush do you use? (Hard-bristled, Medium-bristled, Soft-bristled) (4) Do you use toothpaste to clean your teeth? (5) For those who nominate Toothpaste: Do you use a toothpaste that contains fluoride? Correct responses to all five questions will indicate that EHR patient participants are following the guidance of the ADA on brushing with fluoride toothpaste. These measures will be collected in person via self-report to ascertain baseline status at the beginning of each screening event and scanned directly into the NYU Dentistry remote EHR customized for community outreach events.

- An NYU Dentistry faculty dentist will conduct a head and neck/oral examination on each EHR patient participant and directly enter the results into the customized HER.

- The dentist will review the examination results with the EHR patient participant using a customized walkout statement with screening results and follow-up notes and answer any questions regarding oral health concerns.

Community health workers initial intervention: trained CHWs will deliver a culturally tailored and language-specific oral health promotion program focusing on demonstrations with role playing of proper brushing with fluoride toothpaste and flossing techniques. They will also provide culturally customized literature to the patients.

Acceptability data collection: research staff will conduct a brief exit interview with each interview patient participant regarding acceptability of the intervention and self-efficacy around oral health behaviors (29) using previously validated instruments, requesting permission to contact his/her regular dental provider regarding receipt of a follow-up dental visit and providing a voucher worth $205.00 for oral health care at NYU Dentistry to cover her/his comprehensive oral examination, treatment plan, and prophylaxis at no charge and with no co-payment required. The questions for this survey will be based on our previous oral health promotion program in the Sikh American community to assess acceptability (20). Prior to finalizing the survey for distribution to the interview patient participants, CAB members and research staff will adapt the questions as deemed appropriate for the Chinese American community and to be consistent with the ADA guidance on brushing with fluoride toothpaste.

Feasibility data collection: we will also develop a checklist of 10 key components based on process (e.g., patient engagement) and the curriculum (e.g., topics covered) and ask if each one was covered. Endorsement of 8 of the 10 checklist items (80%) by the non-patient participants will be considered as the bar for success for feasibility. Finally, we will allow for open-ended collection of feedback on the feasibility of each of the partnered program components.

CHW Follow-up Contact with Interview Patient Participants (Feasibility Data Collection)

Community health worker follow-up of oral health care receipt and dental hygiene behaviors will occur at approximately 1 month (window of −7 days to +1 month) after the partnered intervention. This contact may be via telephone or in person.

Feasibility data collection: during this contact, the CHWs will assess whether or not each interview patient participant has received or has scheduled a dental visit. The CHWs will also inquire about use of fluoride toothpaste and frequency of teeth brushing, since the last visit in a modified Oral Health Survey.

If no visit has occurred or been scheduled, the CHW will offer to help schedule a dental visit for the interview patient participant or her/his family members at NYU Dentistry, her/his regular dentist, or one of the project-approved local oral health care providers.

Post-Intervention Activities

- Feasibility and acceptability data collection: semi-structured interviews with participating dentists, CHWs, and other non-patient participants will be conducted.

- Feasibility data collection: EHR data will be reviewed to assess feasibility.

Knowledge Modeling Activities

- Create the interview guide for non-patient participant input into the modeling (pre-intervention).
- Conduct initial non-patient participant interviews to inform model development.
- Conduct the group model-building workshop.
- Conduct follow-up non-patient participant interviews to clarify assumptions.

METHODS: PROCEDURES, EVALUATIONS, AND STATISTICAL CONSIDERATIONS

Workflow Analysis

The team will conduct a workflow analysis, adapted from the Agency for Healthcare Research and Quality recommendations on workflow assessments (55), which includes:

1. Direct entry of patient demographic and appointment (site) information into the EHR by patient service representatives or other authorized users;
2. Ability to scan patient medical history (plus baseline oral health measures of self-reported receipt of a dental visit in the past year and brushing with fluoride toothpaste as recommended) directly into a designated SCREENING folder, followed by transcription and review by faculty dentists;
3. MiFi (Wi-Fi hotspot) configuration to each specific laptop that will only allow EHR access to the matching laptop;
4. Multiple layers of security embedded to comply with all applicable policies;
5. Dual authentication process through Citrix via NYU active directory followed by login to the EHR via NYU ID card;
6. Direct entry of screening results into the EHR by dentists; and
7. Customized walkout statements for interview patient participants with screening results and follow-up notes.

A matrix will be created with three categories: people, documents, and information content. The group adaptation session will be guided by a user-centered design facilitation protocol that sequentially leads the group through presentation of specific remote EHR use cases that include variations on the original EHR data entry screens adapted to the workflow characteristics of local community sites. For each presented use case, the group discussion will focus on the workflow at the site around the role responsibilities (people), documents, and information content (patient medical history, self-reported outcome measures, head and neck/oral examination results) with regards to reviewing customized EHR protocols based on the findings. The discussion will be digitally recorded and each non-patient participant will be given color-coded response sheets to record their perspectives on how to enhance the usefulness of the customized EHR protocols within the community outreach site setting. The digital recordings and response sheets will then be processed, summarized, and converted into adaptation recommendations by the study team.

Next, the study team will transform the recommendations into proposed revisions and document them in revised standard workflow diagrams that build on established workflows to minimize changes at the sites or new work for the dental providers. The insights will then be used to adapt the customized EHR protocols and related workflows, and will be validated in a follow-up group meeting at each site where the adapted EHR screens will be presented and assessed according to the IOF implementation outcome measures of acceptability and adoption. During these follow-up meetings, we will identify any additional workflow variations that the EHR protocols may need to support. Candidate workflows will then be discussed with the project team and other non-patient participants to finalize the adapted workflow integration approach.

Pilot Testing and Live Usability for Remote EHR
In order to account for real world conditions, the intervention will be pilot tested at three Chinese American sites before rolling it out to additional study sites during a planned randomized controlled trial. Early formative observations/short interviews will be conducted with dental teams at each of the pilot sites regarding interaction with the customized EHR. This pilot testing will examine impact on workflow, uncover any new usability problems, and identify any educational needs to be included before large-scale implementation. As providers at the pilot sites engage with the EHR, live-usability testing will be conducted, consisting of direct observations by the research team. Live usability is ideal for observing the impact of new tools on real setting workflows and for observing alternative workflows that can be missed during simulations.

Semi-Structured Interviews Regarding the Remote EHR and Partnered Intervention
Semi-structured interviews will be conducted with participating dentists, CHWs, and other non-patient participants in administrative and technical roles after the intervention. We anticipate conducting approximately 20 interviews before obtaining data saturation. The interviews will be informed by the CFIR and IOF constructs and will assess specific barriers to sustaining the partnered intervention and strategies for addressing those barriers to facilitate integration of the intervention into the routine workflow of the NYU Dentistry Local Community Outreach Programs.

In addition to survey questions, acceptability will also be assessed using open-ended questions, such as:
- Please tell us what you liked most about the program.
- Please tell us what you did not like about the program.
- Please tell us what changes can be made to improve the program.
- Overall program satisfaction (scale of 1–10 answer range).
- How did you feel about the topics covered?
- How did you feel about the length of time of the dental screening and oral health promotion demonstrations?

Feasibility will be assessed among the dentists, CHWs, NYU administrators, and site directors at the three partnering organizations. The following questions will be asked:
- How did you feel about the intake procedures?
- How did you feel about the oral health demonstrations by CHWs?
- How did you feel about the length of time of each dental screening?

Participatory Modeling of Non-Patient Participant Knowledge
The third aim of this study is to model knowledge held by non-patient participants about factors that influence access to oral health care and care-seeking behaviors among low-income, urban Chinese American adults. This information will be used in designing simulation models at multiple levels, from multiple perspectives. A systems science approach will be undertaken to integrate knowledge held by non-patient participants into simulation models to explore alternative paths toward improved health and health care for low-income, urban Chinese American adults via community-based outreach followed by clinical care. These simulation models will be designed using a multi-method approach, in which principles of SD are used to incorporate feedback effects and delays through stocks that accumulate flows (rates of change over time). The SD
approach will be integrated with an ABM framework that is used to appropriately represent dynamics at the community, site, provider, family, and patient levels. The model platform developed for this study will contain multiple model structures that characterize different dynamics and reflect participant input.

### Modeling to Anticipate Effects of Interventions

The models in this platform will simulate implications of hypotheses elicited \textit{a priori} (before implementation of the partnered intervention) from non-patient participants. The \textit{a priori} model platform will enable comparison to models that are later developed with hindsight from implementation of the remote EHR and partnered interventions. This model platform will therefore test the relative effectiveness of the interventions as anticipated under these \textit{a priori} assumptions. Toward this end, this effort will involve design of scoping models that establish a baseline for simulating access at the community level and care-seeking behaviors at the individual level.

A participatory SD modeling process will be undertaken via a group model-building workshop held with non-patient participants as well as semi-structured interviews with individual non-patient participants to elicit targeted model input and feedback on assumptions. The UB Geography Systems Science Modeling Team will work closely with non-patient participants to devise indices for input parameters and indicators for outcomes of simulation experiments. In addition to informing the design of model structures, this participatory approach will enable non-patient participants to better assess the results of the simulation models developed in this \textit{a priori} model platform for authenticity and identification of insights for subsequent implementation research. The resulting model platform will establish a multi-level agent-based GIS framework for simulation modeling of access to oral health care and care-seeking behaviors by low-income, urban Chinese American adults at the community, site, provider, family, and patient levels.

### Study Hypotheses

Our hypotheses for this feasibility and acceptability study are stated next.

#### Primary

Based on exit interviews, patient participants in this study will be satisfied with the partnered intervention components.

#### Secondary

- Based on semi-structured interviews, non-patient participants in the study will be satisfied with the partnered intervention components and remote entry features of EHR software at NYU Dentistry to enter patient information at three Chinese American community sites.
- The knowledge held \textit{a priori} by non-patient participants about factors that influence access to oral health care and care-seeking behaviors among low-income, urban Chinese American adults will enhance the intervention during and/or after the study for further use in future implementations.

### Sample Size Considerations

No formal sample size estimates were performed for this feasibility and acceptability study. The bar for success for both feasibility and acceptability is 80% of enrolled patient and non-patient participants report being satisfied or very satisfied with the partnered intervention components.

### Planned Interim Analyses

Because this is a feasibility and acceptability study, there will be interim reviews of interview data in order to modify aspects of the partnered intervention and the remote EHR processes during the course of the study.

### Final Analysis Plan

Acceptability of the partnered intervention will be assessed through exit interviews of the interview patient participants.

As we did with the Sikh American Families Oral Health Promotion Program (29), we will utilize a pre-post retrospective evaluation design. In this format, all questions will be asked in a single exit interview, but where applicable, will use the phrasing, “Prior to the beginning of the program…” followed by, “At the present time…”. Table 4 provides the measures and definitions of the oral health promotion, self-efficacy, and acceptability measures used in our prior research with the Sikh American community that will be adapted for the present feasibility and acceptability study with the Chinese American community.

The percent change from pre-post will be compared using \textit{t}-tests for proportions. Given that many of the answer choices are non-binary, we will also compare the shift of responses from pre-post across the categories of response using chi-squared tests. The threshold of success for acceptability of the partnered intervention will be that 80% or more of interview patient participants rate all four acceptability questions as agree or strongly agree.

Note that this is not an exhaustive list of questions that will be asked; we are simply highlighting those questions that we will use to quantify acceptability. The full planned questionnaire is attached as an Appendix.

### Health Care Utilization and Oral Health Promotion Measures

Our ultimate health care utilization measure of interest is receipt of a dental visit within the last 12 months. This will be measured using the MEPS definition, where dental visit refers to care by or visits to any type of dental provider. This will allow for direct comparison with Healthy People 2020 Leading Health Indicator OH-7 to increase the proportion of children, adolescents, and adults who used the oral health care system in the past year.

Our central oral health promotion measure is self-reported brushing of teeth for 2 min twice a day with a soft-bristled toothbrush and fluoride toothpaste at the interview patient participant’s 1-month follow-up visit. For the primary health care utilization measure of receipt of a dental visit within the last 12 months, we will also access the NYU Dentistry EHR database and follow-up with oral health care providers identified by participants in HIPAA approved procedures to ascertain receipt of a dental visit in the last 12 months.
Source Documents and Access to Source Data/Documents

Study staff will maintain appropriate research records for this study, in compliance with ICH E6, Section 4.9 and regulatory and institutional requirements for the protection of confidentiality of subjects. Study staff will permit authorized representatives of NIDCR to examine (and when required by applicable law, to copy) research records for the purposes of quality assurance reviews, audits, and evaluation of the study safety, progress, and data validity. Patient participant data will be remotely entered directly into the EHR.

Quality Control and Quality Assurance

The MPIs will be responsible for ensuring that the study is conducted according to the protocol and ensuring data integrity. The MPIs will review the data for safety concerns and data trends at regular intervals, and will promptly report to the IRB and NIDCR any Unanticipated Problem (UP), protocol deviation, or any other significant event that arises during the conduct of the study.

ASSESSMENT OF SAFETY

Specification of Safety Parameters

Safety monitoring for this study will focus on UPs involving risks to subjects, including UPs that meet the definition of a serious adverse event.

Unanticipated Problems

The Office for Human Research Protections (OHRP) considers UPs involving risks to subjects or others to include, in general, any incident, experience, or outcome that meets all of the following criteria: unexpected in terms of nature, severity, or frequency given, (a) the research procedures that are described in the protocol-related documents, such as the IRB-approved research protocol and informed consent document; and (b) the characteristics of the subject population being studied; related or possibly related to participation in the research (“possibly related” means there is a reasonable possibility that the incident, experience, or outcome may have been caused by the procedures involved in the research); and suggests that the research places, subjects, or others at a greater risk of harm (including physical, psychological, economic, or social harm) than was previously known or recognized.

UPS Reporting to IRB and NIDCR

Incidents or events that meet the OHRP criteria for UPs require the creation and completion of an UP report form. OHRP recommends that investigators include the following information when reporting an adverse event, or any other incident, experience, or outcome as an UP to the IRB:

- appropriate identifying information for the research protocol, such as the title, investigator’s name, and the IRB project number;
- a detailed description of the adverse event, incident, experience, or outcome;
- an explanation of the basis for determining that the adverse event, incident, experience, or outcome represents an UP;
- a description of any changes to the protocol or other corrective actions that have been taken or are proposed in response to the UP.

To satisfy the requirement for prompt reporting, UPs will be reported using the following timeline:

- UPs that are serious adverse events will be reported to the IRB and to NIDCR within 1 week of the investigator becoming aware of the event.
- Any other UP will be reported to the IRB and to NIDCR within 2 weeks of the investigator becoming aware of the problem.

### TABLE 4 | Measures and definitions of oral health promotion, self-efficacy, and acceptability measures.

<table>
<thead>
<tr>
<th>Question/domain</th>
<th>Observed change in the Sikh American Program</th>
<th>Expected absolute percent change</th>
<th>Expected relative percent change*</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often do you brush your teeth for at least two minutes?</td>
<td>Increase from 17.9 to 64.7% of those reporting “More than once a day”</td>
<td>+45%</td>
<td>+261%</td>
</tr>
<tr>
<td>How often do you floss?</td>
<td>Increase from 4.4 to 22.1% of those reporting “More than once a day”</td>
<td>+20%</td>
<td>+400%</td>
</tr>
<tr>
<td>How confident…</td>
<td>(a) Increase from 0% to 65.7% of those reporting “Very confident” (b) Increase from 7.4% to 75% of those reporting “Very confident”</td>
<td>(a) +65% (b) +65%</td>
<td>(a) +6,500% (b) +864%</td>
</tr>
<tr>
<td>(a) Do you know how to take good care of your mouth, teeth, and gums? (b) Do you feel asking your dentist or oral hygienist questions?</td>
<td>Agreement with the following statements: (a) Community health worker (CHW) answered my concerns and questions (b) CHW helped me to improve how I take care of my health (c) Information and topics were informative (d) In-person demonstrations helpful in improving oral health</td>
<td>Reported “Strongly agree”</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*We additionally list relative percent change in the case that baseline behaviors among the Chinese American community are very different from those in the Sikh American community.
All UPs should be reported to appropriate institutional officials (as required by an institution’s written reporting procedures), the supporting agency head (or designee), and OHRP within 1 month of the IRB’s receipt of the report of the problem from the investigator.

All UPs will be reported to NIDCR’s centralized reporting system via Rho Product Safety:

Product Safety Fax Line (US): (888) 746-3293.
Product Safety Email: rho_productsafety@rhoworld.com.
General questions about SAE reporting can be directed to the Rho Product Safety Help Line (available 8:00 a.m.–5:00 p.m. Eastern Time):
US: (888) 746-7231.
International: (919) 595-6486.

Halting Rules
This study includes no halting rules.

Study Oversight
The MPIs are responsible for study oversight, in collaboration with the NIDCR Program Official.

ETHICS/DISSEMINATION

Ethical Standard
The MPIs will ensure that this study is conducted in full conformity with the principles set forth in The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research, as drafted by the US National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (April 18, 1979) and codified in 45 CFR Part 46 and/or the ICH E6.

Institutional Review Boards
The protocol, informed consent form(s), and all patient participant and non-patient participant materials were submitted to the IRBs at both NYU Langone Health (study 117-01077) and The State University of New York at Buffalo (study 1749) for review and approval. Approval of both the protocol and the consent forms will be obtained before any patient participant or non-patient participant is enrolled. Any amendment to the protocol will require review and approval by the associated IRBs before the changes are implemented in the study.

Informed Consent Process
Informed consent is a process that is initiated prior to the individual agreeing to participate in the study and continues throughout study participation. Extensive discussion of risks and possible benefits of study participation will be provided to patient participants and their families, if applicable. A consent form describing in detail the study procedures and risks will be given to the patient participant in English or Mandarin Chinese (primary dialect of participating community sites). Consent forms will be IRB-approved, and the patient participant is required to read and review the document or have the document read to him or her. The investigator or designee will explain the research study to the patient participant and answer any questions that may arise. The patient participant will sign the informed consent document prior to any study-related assessments or procedures. Patient participants will be given the opportunity to discuss the study with their surrogates or think about it prior to agreeing to participate. They may withdraw consent at any time throughout the course of the study. A copy of the signed informed consent document will be given to patient participants for their records. The rights and welfare of the patient participants will be protected by emphasizing to them that the quality of their clinical care will not be adversely affected if they decline to participate in this study. The consent process will be documented in the clinical or research record.

Consent for non-patient participant interviews for the knowledge modeling will take place where the interview happens, verbally over the telephone. Dr. Metcalf will send the consent information sheet by email when scheduling the interview, which will be at least 2 days in advance of the interview. To ensuring ongoing consent for follow-up interviews, Dr. Metcalf will remind the non-patient participants about their previous consent and will resend the consent document if needed. Dr. Metcalf will obtain verbal consent for any subsequent recording of information collected during follow-up interviews. Dr. Metcalf will send the consent document ahead of time. Before recording, Dr. Metcalf will ask: have you reviewed the information? Do you have any questions? If the non-patient participant answers “yes” to the first question and “no” to the second question, then she will ask: is it OK if we start the interview now? And if recording the telephone call: is it OK if I begin audio recording? After the participant answers “yes,” the interview will begin.

Women, Minors, and Children
The proposed study will enroll Chinese American adults aged 21 years and older. No children will be included since there are separate and targeted NYU Dentistry programs for this age group. The study population is Chinese American adults living in any of the five boroughs of New York, NY. We estimate that approximately 60% of our enrolled patient participants will be women, based on our experience with conducting community-based screening events. All enrolled patient participants will be of self-reported Chinese ethnicity, to ensure that our partnerships and materials are culturally and linguistically relevant. Study sites will be concentrated in lower Manhattan (Chinatown and the Lower East Side) and the Sunset Park area of Brooklyn, which include dense ethnic enclaves of Chinese Americans.

Subject Confidentiality
Subject confidentiality is strictly held in trust by the MPIs, Co-Is, study staff, and the sponsor(s) and their agents. This confidentiality is extended to cover testing of biological samples and genetic tests in addition to any study information relating to subjects.
The study protocol, documentation, data, and all other information generated will be held in strict confidence. No information concerning the study or the data will be released to any unauthorized third party without prior written approval of the sponsor.

The study monitor or other authorized representatives of the sponsor may inspect all study documents and records required to be maintained by the MPIs, including but not limited to medical records (office, clinic, or hospital) for the study subjects. The modeling site will permit access to such records.

**Certificate of Confidentiality**
To further protect the privacy of study subjects, a Certificate of Confidentiality will be obtained from the US National Institutes of Health (NIH). This certificate protects identifiable research information from forced disclosure. It allows the MPIs and others who have access to research records to refuse to disclose identifying information on research participation in any civil, criminal, administrative, legislative, or other proceeding, whether at the federal, state, or local level. By protecting researchers and institutions from being compelled to disclose information that would identify research subjects, Certificates of Confidentiality help achieve the research objectives and promote participation in studies by helping assure confidentiality and privacy to subjects.

**Data Handling and Record Keeping**
The MPIs are responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported. All source documents will be completed in a neat, legible manner to ensure accurate interpretation of data. The MPIs will maintain adequate case histories of study participants and non-patient participants, including accurate case report forms, and source documentation.

The remote EHR entry of data will be protected by means of a dual authentication process through Citrix via NYU active directory followed by login to the EHR via NYU ID card.

**Data Management Responsibilities**
Data collection and accurate documentation are the responsibility of the study staff under the supervision of the MPIs. All source documents and laboratory reports must be reviewed by the study team and data entry staff, which will ensure that they are accurate and complete. UPs and adverse events must be reviewed by the MPIs or their designees.

**Data Capture Methods**
Patient participant data will be entered remotely into the EHR. Other data will be collected on paper forms and/or digitally recorded.

**Types of Data**
Patient participant data will be captured in the EHR. Data from interviews of patients, research staff, NYU administrators, and providers (dentists and CHWs) will also be captured.

**Study Records Retention**
Study records will be maintained for at least 3 years from the date that the grant federal financial report is submitted to the NIH. Study documents will be retained for a minimum of 2 years after the last approval of a marketing application in an ICH region and until there are no pending or contemplated marketing applications in an ICH region or until at least 2 years have elapsed since the formal discontinuation of clinical development of the investigational product. These documents will be retained for a longer period, however, if required by local regulations. No records will be destroyed without the written consent of the sponsor, if applicable. It is the responsibility of the sponsor to inform the MPIs when these documents no longer need to be retained.

**Protocol Deviations**
A protocol deviation is any noncompliance with the clinical study protocol, Good Clinical Practice, or Manual of Procedures requirements. The noncompliance may be on the part of the subject, the investigator, or study staff. As a result of deviations, corrective actions are to be developed by the study staff and implemented promptly.

These practices are consistent with investigator and sponsor obligations in ICH E6:
- Compliance with Protocol, Sections 4.5.1, 4.5.2, 4.5.3, and 4.5.4.
- Quality Assurance and Quality Control, Section 5.1.1.
- Noncompliance, Sections 5.20.1 and 5.20.2.

All deviations from the protocol must be addressed in study subject source documents and promptly reported to NIDCR and the local IRB, according to their requirements.

**Publication/Data Sharing Policy**
This study will comply with the NIH Public Access Policy, which ensures that the public has access to the published results of NIH funded research. It requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication.

The International Committee of Medical Journal Editors (ICMJE) member journals have adopted a clinical trials registration policy as a condition for publication. The ICMJE defines a clinical trial as any research project that prospectively assigns human subjects to intervention or concurrent comparison or control groups to study the cause-and-effect relationship between a medical intervention and a health outcome. Medical interventions include drugs, surgical procedures, devices, behavioral treatments, process-of-care changes, and the like. Health outcomes include any biomedical or health-related measures obtained in patients or participants, including pharmacokinetic measures and adverse events. The ICMJE policy requires that all clinical trials be registered in a public trials registry, such as ClinicalTrials.gov, which is sponsored by the National Library of Medicine. Other biomedical journals are considering adopting similar policies. For interventional clinical trials performed under NIDCR grants and cooperative agreements, it is the grantee’s responsibility to register the trial in an acceptable registry, so the research results
may be considered for publication in ICMJE member journals. The ICMJE does not review specific studies to determine whether registration is necessary; instead, the committee recommends that researchers who have questions about the need to register error on the side of registration or consult the editorial office of the journal in which they wish to publish.

US Public Law 110–85 (Food and Drug Administration Amendments Act of 2007 or FDAAA), Title VIII, Section 801 mandates that a “responsible party” (i.e., the sponsor or designated principal investigator) register and report results of certain “applicable clinical trials”: Trials of Drugs and Biologics: controlled, clinical investigations, other than Phase I investigations, of a product subject to FDA regulation; and Trials of Devices: controlled trials with health outcomes of a product subject to FDA regulation (other than small feasibility studies) and pediatric postmarket surveillance studies. NIH grantees must take specific steps to ensure compliance with NIH implementation of FDAAA.

Confidentiality

Personal information about potential and enrolled participants will be collected and then de-identified prior to it being shared to ensure confidentiality of participants is maintained before, during, and after the study. Only the MPIs, the biostatistician, and other members of the research team will have access to the final study dataset.

Access to Data

Only investigators and approved researchers added by ethics approval will have access to the final study dataset.

Ancillary and Post-Study Care

The intervention has been developed by NYU Dentistry and a research team with expertise in oral health and health care. We believe that the need to discontinue the intervention will be extremely minimal. If any participant becomes distressed as a result of participation in our study, they will be referred to appropriate counseling support services.

Dissemination Policy

Results from this feasibility and acceptability study will be provided to study participants and disseminated to oral health and health care professionals through presentation at seminars and conferences and publication in scientific journals and relevant media. We will adhere to all guidelines for authorship.

Appendices

The patient consent brochure, non-patient consent brochure, and consent signature page will be provided to participants and their authorized surrogates and are available as Appendices.

STRENGTHS AND LIMITATIONS

The strengths of this feasibility and acceptability study include the expertise and experience of the involved researchers, providers, and administrators, and the commitment of NYU Dentistry to improve its Local Community Outreach Programs to meet the needs of the local Chinese American community. Further, the Center for the Study of Asian American Health at NYU Medicine is the only center of its kind in the United States solely dedicated to research and evaluation on Asian American health and health disparities. Thus, this study holds the potential to fill a care gap for this diverse and growing population. In particular, Chinese are the largest Asian ethnic group in New York, NY, with higher poverty rates for working age and older adults relative to all residents (56). As NYU Dentistry and NYU Medicine have partnered on CBPR initiatives using CHW models (21, 29), there is confidence in our ability to adapt materials and programming for this new population/setting of Chinese American outreach sites. Further, in longstanding collaboration with The State University of New York at Buffalo, our research team has examined how factors at multiple levels contribute to oral health and care-seeking behaviors of racial/ethnic minority older adults (37, 38). Leveraging our portfolio of systems science models and group model-building expertise and experience, we plan to engage with our partners to understand the dynamic complexity of our interventions and simulate alternate scenarios, in concert with a recent NIH funding opportunity announcement (57). Finally, integrating multiple scientific approaches (implementation science, CBPR, and systems science) and utilizing remote EHR capabilities to enhance patient care and tracking are notable strengths of this study.

This feasibility and acceptability study also has certain limitations. First, findings may not be generalizable to other settings and locales. Nonetheless, by furnishing our protocol to the research community, other implementation scientists may adapt it for their local needs. Second, the study is only funded for 1 year. Thus, we will not be able to track patient participants screened during the second half of the funding period to determine whether or not they visited a dental provide in the complete follow-up year, as per Leading Health Indicator OH-7 (6). Still, this funding provides an opportunity to ascertain the feasibility and acceptability of our partnered intervention, and strengthen our methods toward designing and implementing a randomized controlled trial of a participatory, multi-level, partnered intervention to improve the oral and general health of low-income Chinese American adults. Further, while the pre-post retrospective evaluation design to assess the acceptability of the partnered intervention is considered a practical method to evaluate learning from an educational program (58), it has certain limitations. In particular, recall bias may affect the quality of the data collected due to potentially inaccurate or skewed memories of patients regarding their prior attitudes, emotions, behaviors, and experiences. Finally, social desirability bias may affect patients’ self-reported brushing of teeth for 2 min twice a day with a soft-bristled toothbrush and fluoride toothpaste.

ETHICS STATEMENT

This feasibility and acceptability study will be conducted in accordance with the International Council on Harmonization guidelines for Good Clinical Practice (ICH E6), the Code of Federal Regulations on the Protection of Human Subjects (45 CFR Part 46), and the NIDCR Clinical Terms of Award. All subjects gave written informed consent in accordance with the
Declaration of Helsinki. The protocol was approved by the IRBs at both NYU Langone Health (study i17-01077) and The State University of New York at Buffalo (study 1749).

AUTHOR CONTRIBUTIONS

MN conceived the study and led the writing. SM and QZ contributed systems science and geographic expertise. SY and CT-S contributed CBPR and evaluation expertise. XG contributed health and disease science and geographic expertise. All authors contributed to the drafting and approved the final version of the paper.

ACKNOWLEDGMENTS

Members of the Implementing a Participatory, Multi-level Intervention to Improve Asian American Health Research Team contributed to the conceptualization, administration, grant writing, direction, and/or implementation of this study and are not already named as authors include (in alphabetical order): Nadia Islam, Smiti Nadkami, Janet Pan, Rebecca Park, Sharon Perelman, Andrew Schenkel, Andrea Troxel, Mark Wolff, and Jennifer Zanowiai. The authors are grateful to the NIDCR and Rho, Inc., scientists who contributed expertise and support to the authors in writing of protocol over a 3-month period of intense collaboration.

FUNDING

Funding for this feasibility and acceptability study is provided by the NIDCR (1U56DE027447-01) for the project, Implementing a Participatory, Multi-level Intervention to Improve Asian American Health (MPIs: Northridge, Trinh-Shevrin, Metcalf).

REFERENCES


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
A Mixed Methods Approach to Evaluate Partnerships and Implementation of the Massachusetts Prevention and Wellness Trust Fund

Rebekka M. Lee1,2, Shoba Ramanadhan1,3, Gina R. Kruse1,4 and Charles Deutsch1

1 Clinical and Translational Science Center, Harvard Medical School, Boston, MA, United States; 2 Prevention Research Center, Harvard T.H. Chan School of Public Health, Boston, MA, United States; 3 Center for Community-Based Research, Dana Farber Cancer Institute, Boston, MA, United States; 4 Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA, United States

Background: Strong partnerships are critical to integrate evidence-based prevention interventions within clinical and community-based settings, offering multilevel and sustainable solutions to complex health issues. As part of Massachusetts' 2012 health reform, The Prevention and Wellness Trust Fund (PWTF) funded nine local partnerships throughout the state to address hypertension, pediatric asthma, falls among older adults, and tobacco use. The initiative was designed to improve health outcomes through prevention and disease management strategies and reduce healthcare costs.

Purpose: Describe the mixed-methods study design for investigating PWTF implementation.

Methods: The Consolidated Framework for Implementation Research guided the development of this evaluation. First, the study team conducted semi-structured qualitative interviews with leaders from each of nine partnerships to document partnership development and function, intervention adaptation and delivery, and the influence of contextual factors on implementation. The interview findings were used to develop a quantitative survey to assess the implementation experiences of 172 staff from clinical and community-based settings and a social network analysis to assess changes in the relationships among 72 PWTF partner organizations. The quantitative survey data on ratings of perceived implementation success were used to purposively select 24 staff for interviews to explore the most successful experiences of implementing evidence-based interventions for each of the four conditions.

Conclusions: This mixed-methods approach for evaluation of implementation of evidence-based prevention interventions by PWTF partnerships can help decision-makers set future priorities for implementing and assessing clinical-community partnerships focused on prevention.

Keywords: implementation science, mixed methods research, asthma, hypertension, falls, tobacco
INTRODUCTION

The delivery of preventive services in community-based and clinical settings has tremendous potential to improve population health. However, these community and clinic-based preventive activities are rarely coordinated (1), even with evidence that clinical-community partnerships can improve health outcomes including smoking abstinence, perceived physical health, cholesterol levels and hypertension (2, 3). The potential of community-clinical partnerships to improve health is further emphasized by the finding that neighborhood or community-level determinants of health also impact the way patients interact with the healthcare system as measured by hospital readmissions (4) and emergency room visits (5). As healthcare systems become increasingly accountable for improving the health of populations, strategies for linking clinical systems and community-based partners are becoming essential (6).

Clinical-community collaborations offer an opportunity to create multi-level, sustainable change. Thousands of coalitions, alliances, and other forms of inter-organizational health focused partnerships were formed over the past two decades (7–9). These intersectoral partnerships are critical for addressing complex public health challenges. They can marshal complementary human and social capital, embed interventions in the broader public health system, and offer opportunities to address problems that cannot be solved by an organization or sector in isolation (8–11). Although collaboration across sectors or institution types is not without its challenges (12), coalitions and intersectoral partnerships have successfully impacted health disparities broadly (13), as well as in improved diabetes, HIV/AIDS, and substance abuse outcomes (14–16).

The clinical-community partnerships in this project implemented evidence-based interventions that address hypertension, pediatric asthma, falls among older adults, and tobacco use throughout Massachusetts. In 2012, as the second stage in Massachusetts’ ground-breaking health reform initiative, the legislature passed Massachusetts General Law Chapter 224 (17). Among other things, it established the Prevention and Wellness Trust Fund (PWTF), which provided more than $42 million over 4 years to nine community-clinical partnerships. The Massachusetts Department of Public Health led the initiative, competitively selecting nine partnerships in diverse communities across the state and providing technical assistance to implement specified evidence-based interventions. The conditions and interventions were chosen for implementation because they were determined to be more likely than others to show changes in outcomes and costs, and positive return on investment, in the span of 3 years. The nine chosen communities exceeded state-wide prevalence of the priority conditions, were more racially and ethnically mixed, and had higher rates of poverty than the state average (18). The funded partnerships varied in configuration and ranged in size from 40,000 to 140,000 people; some were single cities, others included multiple cities and towns, and one constituted an entire county. Fifteen percent of the state population resides within the nine funded partnerships. All partnerships included a city/regional planning agency, a clinical health provider, and a community-based organization. Their size range from 6 to 15 participating organizations. More details on the PWTF partnerships, decisions, interventions, and model are available in the project final report (19).

The initiative began in 2014 with a 6–9 month planning stage focused on capacity building. Communities developed partnerships among clinical providers and community-based organizations that linked and coordinated clinical and community-based strategies. The request for response specified that at least one intervention must involve bi-directional referrals from clinical to community organizations with feedback loops. For example, a community health center might partner with the YMCA to develop a system in which patients screened as hypertensive or at risk for falls are referred to community programming, and conversely YMCA members who express needs for clinical services are referred to the community health center. For most of the partnerships, full implementation began early in 2015. Table 1 lists the clinical and community evidence-based interventions for each health condition. Of the nine partnerships, all selected hypertension, eight selected falls among older adults, five selected tobacco cessation, and six chose pediatric asthma. MDPH provided grantee support, such as individualized technical assistance in evidence-based interventions, learning sessions to facilitate knowledge development and sharing across all grantees, and quality improvement evaluation. Partnerships were encouraged to culturally adapt interventions to meet the needs of their local communities.

The communities were required to jointly fund a rigorous independent evaluation of the PWTF to determine if it met its explicit legislative objectives: (1) a reduction in the prevalence of preventable health conditions; (2) a reduction in health care costs or the growth in health care cost trends associated with these conditions; and (3) an assessment of which populations benefited from any reduction. While not specified in the authorizing legislation, the Prevention and Wellness Advisory Board (PWAB) created by Chapter 224 strongly recommended the additional systematic collection of data that illustrate the implementation experiences in PWTF communities.

The purpose of this paper is to present a mixed methods approach to assess the PWTF implementation experience. While an outcome evaluation is critical to establishing success, embedding quantitative surveys and qualitative interviews that assess how these partnerships function and what contextual factors influence implementation will help to provide actionable findings. This paper draws upon implementation science, social network analysis, and a mixed methods design to understand these complexities.

First, the field of dissemination and implementation science is concerned with generating knowledge beyond clinical trials and effectiveness research to investigate change in real-world settings. In this study, we define implementation “as the way and degree to which an intervention is put into place in a given setting” (20). Fundamental to implementation science is the concept of integrating evidence-based interventions within a community or clinical setting and creating partnerships and...
TABLE 1 | Clinical and community interventions implemented as part of the Prevention and Wellness Trust Fund.

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Tier</th>
<th>Partnerships</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HYPERTENSION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evidence-Based Guidelines for Hypertension Screening</td>
<td>Clinical 1</td>
<td>9</td>
</tr>
<tr>
<td>Chronic Disease Self-Management Programs</td>
<td>Community 1</td>
<td>8</td>
</tr>
<tr>
<td>Self-Monitored Blood Pressure with Additional Support</td>
<td>Community 2</td>
<td>6</td>
</tr>
<tr>
<td>Diabetes Prevention Program for Patients with Hypertension and Pre-Diabetes</td>
<td>Community 2</td>
<td>3</td>
</tr>
<tr>
<td><strong>FALLS AMONG OLDER ADULTS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stopping Elderly Deaths, Accidents, and Injuries: Clinical Risk Assessment</td>
<td>Clinical 1</td>
<td>8</td>
</tr>
<tr>
<td>Assisted Home Safety Assessment</td>
<td>Community 2</td>
<td>8</td>
</tr>
<tr>
<td>Matter of Balance</td>
<td>Community 2</td>
<td>8</td>
</tr>
<tr>
<td>Tai Chi</td>
<td>Community 2</td>
<td>6</td>
</tr>
<tr>
<td><strong>TOBACCO CESSATION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Preventive Services Task Force Screening Guidelines</td>
<td>Clinical 1</td>
<td>5</td>
</tr>
<tr>
<td>Tobacco Cessation Counseling</td>
<td>Community 1</td>
<td>5</td>
</tr>
<tr>
<td>Promoting Smoke-Free Environments</td>
<td>Community 2</td>
<td>5</td>
</tr>
<tr>
<td><strong>PEdiATRIC aSTMa</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care Management for High Risk Asthma Patients</td>
<td>Clinical 1</td>
<td>6</td>
</tr>
<tr>
<td>Asthma Self-Management in Primary Care</td>
<td>Clinical 2</td>
<td>4</td>
</tr>
<tr>
<td>Home-Based Multi-Trigger Multi-Component Intervention</td>
<td>Community 1</td>
<td>5</td>
</tr>
<tr>
<td>Comprehensive School-Based Asthma Programs</td>
<td>Community 2</td>
<td>4</td>
</tr>
<tr>
<td>Comprehensive Head Start-Based Asthma Programs</td>
<td>Community 2</td>
<td>2</td>
</tr>
</tbody>
</table>

supportive delivery systems to support the use of evidence-based interventions. At the core of this science is inquiry into the contextual factors that influence successful implementation of evidence-based interventions. To ground our inquiry, we applied the Consolidated Framework for Implementation Research (CFIR), an established framework that supports identification of actionable factors that influence success within five domains: the inner setting, the outer settings, characteristics of individuals, characteristics of the intervention, and processes (21).

Next, it was important to examine the composition, structure, and functions of the PWTF partnerships in the context of implementing evidence-based interventions. Social network analysis is a natural fit for evaluation of the function and impact of community-clinical partnerships, as it focuses on relationships (here, between organizations) and takes a systems perspective (22). Social network analysis has been applied effectively to the study of a range of collaborative efforts among organizations engaged in health promotion activities (23–25).

Using the methods of social network analysis, it becomes possible to assess the form and function of a network, identify key actors and the types of resources exchanged across the network, assess the sustainability and strength of relationships, assess opportunities to strengthen the network’s impact on a set of health outcomes, and assess challenges or drawbacks to collaboration (11). In this way, social network analysis affords the opportunity to explore the ways in which community-clinical partnership networks can be utilized to create change and achieve intended implementation outcomes (and ultimately, intended health outcomes) in the organizations and communities of interest.

Finally, mixed methods research is the collection and analysis of quantitative and qualitative data, which is often employed to understand complex research problems for which one methodology is not sufficient (26). Mixed methods studies must use rigorous quantitative and qualitative methods and explicitly integrate or link these two types of data for a more comprehensive investigation of the topic at hand (26). Using mixed methods can be helpful for understanding the perceptions of practitioners and end-users of a given evidence-based intervention (27). A mixed methods design also aligns well with the need to conduct multi-level assessments of implementation efforts (e.g., collecting data at the community, clinic, provider, and patient levels) (28, 29). In this study, we use a multi-phase, explanatory sequential mixed methods design embedded in a large evaluation project to gain a more comprehensive understanding of implementation of the Prevention and Wellness Trust Fund interventions (Figure 1) (26, 30). Building three rapid phases of data collection and analysis upon one another is intended to explain what success looks like in this state-wide implementation of clinical-community linkages to build population-level disease prevention and management systems.

This mixed methods external evaluation will be useful to a variety of stakeholders, including legislators and other policymakers who need to know what PWTF accomplished and what next steps are indicated; implementing communities and agencies who need to know what worked and what didn’t, and for whom; and other communities that want to learn from the PWTF experience.

MATERIALS AND METHODS

We used a multi-phase explanatory mixed methods design embedded in a larger evaluation to investigate what interventions work for whom and in what settings—key issues at the core of implementation science (see Figure 1). First, we conducted semi-structured qualitative telephone interviews (lasting about 1.5 h) with at least two leaders from each of the nine partnerships. Key informant interviews are in-depth discussions that offer insight into participants’ perceptions and opinions and are suited for exploratory research (31). They are often conducted with an individual, but we chose to conduct them with leadership teams to gather high-level perspective and a sense of daily implementation efforts. The interview findings were used to develop a quantitative survey to assess the implementation experiences of 172 staff from participating clinical and community-based organizations and a social network analysis to assess changes in the relationships among 70 PWTF organizations. The quantitative survey data on ratings of perceived implementation success were used to
purposively select 24 staff for interviews. These 1.5-h interviews (in person whenever possible) were intended to explore the most successful experiences of implementing evidence-based hypertension, falls, tobacco, and asthma interventions. We chose interviews at this stage rather than staff focus groups because we sampled different cadres of staff (e.g., physicians, partnership coordinators, community health workers). We expected some staff would be more comfortable describing challenges or barriers to implementation in one-on-one interviews versus focus groups which may have included more senior staff and leaders from their communities. Detailed descriptions of each of the phases of the mixed methods implementation evaluation are below and described visually in Figures 1, 2 with details on the project timeline, data collection and analyses activities, and products. The Consolidated Framework for Implementation Research (CFIR) guided the development of this evaluation (21). The Harvard Office of Human Research Administration (IRB) determined that full review and approval was not required for this study. It has been approved by the Office of Human Research Administration staff and the proposal was reviewed by the Department of Public Health’s Institutional Review Board.

Phase 1: Qualitative Interviews With Coordinating Partners

In March 2016, key informant interviews in Phase 1 served as an initial, high-level qualitative exploration of the implementation experience in each partnership and helped to adapt existing survey items to identify contextual influences on PWTF implementation in Phase 2.

Sampling, Recruitment, and Administration

Each partnership had one organization that served as the coordinating partner, meaning that it was responsible for leading and managing the initiative. The Massachusetts Department of Public Health identified participants from the coordinating partners for the Phase 1 qualitative interviews. The 2–4 key informants from each community included the current PWTF project manager from each partnership, plus additional interviewees with a large breadth of knowledge about this project. Participants included health department directors, community health center senior leadership, healthcare system administrators, and past project managers in communities that had experienced leadership turnover. Prior to interviews, the
study team emailed each PWTF project manager a one-page overview detailing the purpose and expectations of each phase of the implementation evaluation. All interviews were scheduled via email and conducted over the phone at the convenience of coordinating partners. The research team conducted 1.5-h telephone interviews with each coordinating partner team. All coordinating partners agreed to participate in Phase 1 interviews.

Measures
Implementation constructs explored in the Phase 1 interview included the implementation experience as well as an exploration of the contextual influences on implementation. To capture implementation experience, we included prompts related to buy-in among leadership and staff, details of intervention adaptation and delivery, the role of community health workers in supporting community-clinical partnerships to implement evidence-based interventions, and the connection between intervention implementation and health equity issues. The research team adapted an existing interview guide (32) based on the Consolidated Framework for Implementation Research (CFIR) to the PWTF settings and outcomes, attending to each of the five CFIR domains: inner setting (e.g., leadership engagement, resources) characteristics of the intervention (e.g., complexity, relative advantage), characteristics of individuals (e.g., role, turnover), outer setting (e.g., community context), and processes (e.g., planning, engaging champions) (21). The full interview guide is available in Supplementary Material 1 and example of qualitative interview questions appear in Table 2.

The social network analysis portion of the interview guide examined two classes of networks: (a) intra-partnership networks (relationships between PWTF organizations within each of the nine partnerships) and (b) inter-partnership networks (relationships between the nine partnerships). For the intra-partnership network assessment, the first step was to define the set of organizations of interest; in this case, all organizations involved with PWTF implementation (33). For each partnership we used the list from the MDPH as a starting point and then reviewed it with partners to revise as needed. Second, the interview guide included prompts to define relationships of interest. The literature suggests that important relationships linked to creating practice change in healthcare settings include communication, collaboration or competition, exertion of influence, and exchanging resources (25, 34). We asked about these and also prompted respondents to identify other important interactions or exchanges that supported their PWTF goals. Finally, we asked a set of questions to explore the role.
of additional, unofficial partners in the PWTF initiative. For example, a given community-based organization may be the official delivery site for a given evidence-based intervention, but may link with other local organizations for recruitment or other activities.

For the inter-partnership assessment, the interview guide focused on relationships among the nine participating partnerships, as they had been brought together as part of a quality improvement learning collaborative to support PWTF goals. The interviews focused on the range of network relationships involved in implementing evidence-based interventions through the PWTF. We also asked about the range of benefits derived from engaging with other partnerships and expected sustainability of these relationships.

Data Management and Analysis
Interview recordings were transcribed verbatim. Data were managed and prepared for analysis using NVivo qualitative data analysis software Version 11 (QSR International Pty Ltd. 2012. Melbourne, Australia). The research team reviewed transcripts for key constructs to include in the Phase 2 quantitative implementation and social network surveys. We conducted a cross-case analysis that began deductively coding according to contextual factors from CFIR, and then inductively added codes for new patterns and themes. Rigor was ensured with analysis triangulation; all interviews were coded by two researchers to ensure multiple perspectives (35, 36). Interview data were integrated with the phase 2 survey and phase 3 interview data, looking for concordant and discordant results (26).

Phase 2: Quantitative Surveys
During May and June 2016 in Phase 2 of this evaluation, we fielded two online surveys to quantitatively identify the contextual factors that influenced implementation of the evidence-based interventions and assess the social networks within and between each partnership. Both surveys helped to adapt an existing guide (32) for follow-up in-depth interviews in Phase 3.

Sampling, Recruitment, and Administration
The research team worked with the Massachusetts Department of Public Health and coordinating partners to generate a list of all organizations that were part of each partnership. Next, coordinating partners indicated the health conditions and evidence-based interventions associated with each organization and listed the names, roles/titles, and email addresses for 1–3 contacts at each organization who were involved with implementing the evidence-based interventions. They were asked to include clinical staff of varying levels (doctors, nurses, and medical assistants), practitioners in community-based settings, and community health workers. One week prior to launching the surveys, the study team emailed each PWTF project manager to disseminate a one-page overview detailing the broad content...
areas of focus on the survey. Project managers from each partnership shared the overview with participants.

Both surveys were conducted online via REDCap electronic data capture tools (37). The implementation survey was administered to all contacts identified by the coordinating partners (N = 214). The social network survey was administered to one representative at each organization designated as the lead for the PWTF (N = 90). Participants were invited to complete the surveys by email. They were given a 2-week window to respond to the surveys, with reminders sent at 1 week and 1 day before the official close. Coordinating partners assisted in encouraging survey participation. Participants were incentivized to complete the implementation survey with a chance to win a raffle for a $75 gift card. A total of 172 individuals completed the implementation survey (response rate = 80%) and 72 people completed the social network survey (response rate = 80%).

Measures
The research team adapted existing validated survey items (38, 39) to the PWTF settings and outcomes using findings gleaned from the Phase 1 interviews. Items assessed the perceived degree of implementation for each evidence-based intervention as well as contextual domains in the CFIR (21). A 4-point Likert scale captured the degree of implementation, with the following ratings: 0 (no implementation); 1 (“we are in the early stages of implementation”); 2 (“we have implemented this strategy, but inconsistently”); and 3 (“we have implemented this intervention fully and systematically”). The CFIR survey items were measured on a 5-point Likert scale with responses ranging from 1—strongly disagree to 5—strongly agree. We also included items to capture title, role, age, gender, race/ethnicity, education, language spoken, and years of experience. Adaptations to the survey were made based on qualitative data provided by the coordinating partners in Phase 1. For instance, sufficient staffing and data systems/IT support were frequently named as important resources influencing implementation; therefore, we created discrete items to assess these factors quantitatively on the survey. Using the qualitative data to adapt the quantitative survey ensured we could measure the frequency of these contextual influences in the large pool of 172 clinical and community-based implementers. The full survey is available in Supplementary Material 2 and Table 2 includes examples of survey items.

The quantitative, intra-partnership social network analysis utilized the list of organizations involved with PWTF implementation from the Phase 1 interviews and asked about relationships with all other members of the partnership. For example, if a given partnership included 7 organizations, we surveyed each organization about their relationships with the other 6 organizations. The social network analysis focused on a core set of relationships identified in Phase 1 as important for implementation: collaboration, sharing information/resources, sending referrals, receiving referrals, providing/receiving technical assistance or capacity-building, providing/receiving access to community members. We also asked questions about the sustainability of reported connections after funding is completed. Finally, we asked questions to prompt respondents to identify up to five additional partners involved in the execution of the evidence-based program or strategy. The quantitative, inter-partnership social network analysis included questions about relationships (using the same list provided above) with the other partnerships. Once more we asked about expected sustainability of connections after funding ends.

Data Management and Analysis
The research team analyzed quantitative survey data in SAS v9.4 (SAS Institute: Cary, NC). We calculated descriptive statistics (e.g., means of implementation outcomes and CFIR constructs) for all outcomes. A summary score for each evidence-based intervention was created for each partnership by averaging ratings of implementation from all respondents in each partnership. These 4-point scale summary scores were used to classify partnerships as “high implementation” using self-reported scores for each health condition. High implementation partnerships for each condition had summary scores for each evidence-based intervention that were higher than the PWTF average. Social network data were analyzed using a combination of the dedicated network analysis software UCINET (Analytic Technologies: Lexington, KY) and SAS v9.4. Quantitative social network analyses emphasized analysis of the relationships within the official set of network members for each partnership. The analyses linked social network metrics with implementation outcomes.

Phase 3: Qualitative Interviews With Implementers
In July and August 2016, the final phase of our evaluation, we conducted follow-up in-depth interviews with practitioners charged with implementation. The interviews focused on developing a more comprehensive understanding of the experience of implementing the evidence-based hypertension, falls, asthma, and tobacco interventions in real world clinical and community settings.

Sampling, Recruitment, and Administration
The research team sampled “high implementation” partnerships for participation in the Phase 3 interviews. The 4-point summary scores from Phase 2 surveys were used to classify partnerships as “high implementation” using self-reported scores for each health condition.

After high implementation partnerships were identified, the research team sampled 4–6 individuals (at least one clinical partner and one community partner) from each partnership for interviews. These individuals were purposively sampled from the list of implementation survey respondents in an effort to conduct information-rich interviews. For instance, Phase 3 interviews for falls among older adults in one partnership included speaking with a community health worker who conducted falls assessments and referrals within a community health center, a falls prevention coordinator from an elder services organization responsible for home safety assessments, folks leading Matter of Balance and Tai Chi classes at the YMCA and via city recreation, as well as the director of a local non-profit organization.
All 1.5-h interviews were scheduled via email and conducted in-person at the convenience of the participants whenever possible (two interviews were conducted over the phone). Interviews were audio recorded and transcribed verbatim. Participants were compensated with a $25 gift card. All people invited for Phase 3 interviews agreed to participate.

Measures
Similar to the Phase 1 formative interviews, the research team adapted an existing interview guide (32) based on the CFIR to the PWTF settings and outcomes. The adaptation included tailoring the interview to investigate findings from the quantitative surveys of Phase 2. Targeted probes for CFIR items with the highest or lowest average ratings on the survey were added to the interview. This was done to explore barriers and facilitators to implementation in greater depth. For example, respondents’ extreme rating of the complexity of interventions and resources such as staffing led our team to add probes to the interview guide to gain a better understanding of what intervention complexity and staffing constraints looked like from the perspectives of those who were implementing the interventions in real world settings. Implementation constructs explored in the Phase 3 follow-up interview included the experience of implementing specific evidence-based interventions and an exploration of the contextual influences on implementation. Elements of the implementation experience include buy-in among leadership and staff, a description of how interventions were adapted and delivered, the role of community health workers, and strategies to address health equity. Clinical partners were also asked to discuss how quality of care initiatives impacted implementation of the PWTF interventions (40). All five CFIR domains were explored in this phase for each target health condition (21). The full interview guide appears in Supplementary Material 3 and there are examples of qualitative interview items in Table 2.

The analysis of Phase 2 network data highlighted the diversity of partnership structure for organizations working together to implement evidence-based interventions through the PWTF. We explored this further by asking implementers to describe their experiences with community-clinical linkages as part of the PWTF initiative. We also asked a series of questions about partnership sustainability to compare and contrast descriptions provided by implementers vs. descriptions provided by partnership leaders (Phase 1).

Data Management and Analysis
All interview audio-recordings were transcribed. Data were managed and prepared for analysis using NVivo qualitative data analysis software Version 11 (QSR International Pty Ltd. 2012. Melbourne, Australia). We conducted a cross-case analysis that began deductively coding according to contextual factors from CFIR, and then inductively added codes for new patterns and themes (35, 36). One-third (8 of 24) of transcripts were coded by a second researcher to build consensus around all codes and themes. Phase 3 interview data were integrated with survey and Phase 1 key informant interview data, looking for concordant and discordant results (26).

DISCUSSION
This paper describes the design of a mixed methods approach for evaluating the implementation of clinical-community partnerships through The Prevention and Wellness Trust Fund. This study design will help us gain a comprehensive understanding of this complex approach for engaging communities in implementing evidence-based interventions across Massachusetts. To create an evaluation protocol that was truly mixed-methods, rather than simply multi-method, it was critical to explicitly and strategically find points in the evaluation process to integrate our qualitative data (41). In our multi-phase, explanatory sequential mixed methods design embedded in the larger PWTF evaluation, data were integrated or linked in several ways. First, while the initial mandated evaluation focused solely on the analysis of large quantitative datasets of medical claims, hospital discharges, and aggregated electronic health records, the PWTF advisory board and our research study team also prioritized embedding qualitative data into the larger evaluation to understand the complexities of the local implementation experiences. We also integrated quantitative and qualitative data to build implementation survey measures. The initial interviews with key informants were used to prioritize and adapt survey items for a tailored quantitative assessment of partnership social networks and implementation of the PWTF evidence-based interventions with a broader sample of implementation stakeholders in phase 2. Additionally, the study followed up on surveys with a second round of interviews as a means of explaining the quantitative results in greater depth. In this explanatory process, we used quantitative data on perceived level of implementation to sample “high implementation” partnerships and create qualitative probes to examine contextual implementation factors that were quantitatively rated as influential. This complex design presented the challenge of multiple phases depending on the success of earlier phases and determining how much data is sufficient to move forward to each subsequent mixed methods phase. For example, deciding how much quantitative analysis of the online survey should be conducted to inform the sampling and adaptation of the qualitative follow-up interviews.

The Prevention and Wellness Trust Fund sought to build and use partnerships to implement complex interventions in complex systems (42), meaning that a group of connected, “un-siloed” interventions addressing four priority health conditions were implemented in coordination across a variety of settings (e.g., hospitals, community health centers, schools, YMCAs, housing). By measuring the function and impact of partnerships within and between communities implementing evidence-based prevention programs, this evaluation is designed to better understand how to set up and support community-based prevention efforts. Accountable Care Organizations, which strive to develop clinic-community partnerships to improve the health of populations may use PWTF as a prototype. Using implementation science, interviews and surveys may help identify best practices for tailoring evidence-based interventions to unique contexts and constituents. The mixed methods study design also allows us to detail the challenges of clinical-community linkages, which
are vital both in the narrow sense of promoting the use of specific evidence-based programs or practices, but also in a broader sense of, supporting sustainable community-level, systems changes (43).

The use of a mixed methods approach to understanding the implementation of evidence-based practices in clinical-community partnerships draws on the strengths of both qualitative and quantitative methods, but it is not without limitations. First, time constraints presented challenges in several ways, given that the external evaluation was only funded for the second year of a three-year implementation period. Limited time meant that our study was only able to conduct in-depth follow-up interviews with people implementing the interventions in “high implementation” partnerships. If we had more time, we could have prioritized exploring the implementation process and contextual factors within partnerships that have less success in greater depth with follow-up interviews that could further our understanding of implementation challenges. Time also limited our ability to use more objective quantitative measures, such as program reach or changes in clinical outcomes, to sample “high implementation” partnerships. We were also limited in our ability to evaluate how partnerships were trained and subsequently implemented interventions to address health equity, with only one question on interviews directed toward this topic.

In sum, this paper details the research protocol for the external evaluation of the implementation of the Prevention and Wellness Trust Fund. Subsequent implementation research from this project aims to describe how the hypertension, falls, asthma, and tobacco evidence-based interventions were implemented and identify actionable contextual factors that influenced implementation in the nine partnerships. The mixed methods approach will provide data that appeals to influenced implementation in the nine partnerships. The findings from this study will be valuable for understanding what PWTF has accomplished and to help other communities planning to set-up or support community-clinical partnerships to deliver evidence-based preventive services.

**AUTHOR CONTRIBUTIONS**

RL served as the lead author on the paper, contributing to conceptualization, literature summary, development of data collection measures, drafting/editing all sections of the paper, tables, and figures. SR contributed to conceptualization of the paper, literature summary, development of data collection measures, and drafting/editing methods and discussion. GK contributed to conceptualization of the paper, development of data collection measures, and drafting/editing the introduction and methods. CD served as senior author on the paper, contributing to conceptualization of the paper, development of data collection measures, and drafting/editing the introduction and discussion.

**FUNDING**

The study was funded by the Commonwealth of Massachusetts/Department of Public Health (INTF4250HH2500224018), NIH/NIMH (2002362059), and NIH/NCATS (4UL1TR001102-04). The views and opinions in this publication do not necessarily reflect the views and opinions of the Massachusetts Department of Public Health.

**ACKNOWLEDGMENTS**

In addition to our funder that provided valuable details on the PWTF model detailed in the introduction, we would like to thank the almost 200 practitioners and leaders throughout Massachusetts who shared their perspectives on implementation of PWTF through interviews and surveys as well as James Daly, Amy Cantor, and Queen Alike who helped to conduct interviews and analyze data for the project.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2018.00150/full#supplementary-material

**REFERENCES**


10. Provan KG, Nakama L, Veazie MA, Teufel-Shone NI, Huddleston C. Building community capacity around chronic disease services through
Goerzen A, Beamish PW. The effect of alliance network diversity on multina
10.1002/smj.447

Hennessey LS, Smith ML, Esperza AA, Hrushow A, Moore M, Reed DF. The community action model: a collaborative-driven model designed
10.2105/AJPH.2004.077004


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Lee, Ramanadhan, Kruse and Deutsch. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Developing a Survey Tool to Assess Implementation of Evidence-Based Chronic Disease Prevention in Public Health Settings Across Four Countries

Elizabeth L. Budd 1*, Xiangji Ying 2, Katherine A. Stamatakis 3, Anna J. deRuyter 2, Zhaoxin Wang 4, Pauline Sung 5, Tahna Pettman 6, Rebecca Armstrong 6, Rodrigo Reis 2 7 and Ross C. Brownson 2

1 Prevention Science Institute, College of Education, University of Oregon, Eugene, OR, United States, 2 Prevention Research Center, Brown School, Washington University in St. Louis, St. Louis, MO, United States, 3 College for Public Health and Social Justice, St. Louis University, St. Louis, MO, United States, 4 Tongji University School of Medicine, Shanghai, China, 5 Department of Applied Social Sciences, The Hong Kong Polytechnic University, Kowloon, China, 6 Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia, 7 School of Health and Biosciences, Pontifical Catholic University of Parana, Curitiba, Brazil

Background: Understanding the contextual factors that influence the dissemination and implementation of evidence-based chronic disease prevention (EBCDP) interventions in public health settings across countries could inform strategies to support the dissemination and implementation of EBCDP interventions globally and more effectively prevent chronic diseases. A survey tool to use across diverse countries is lacking. This study describes the development and reliability testing of a survey tool to assess the stage of dissemination, multi-level contextual factors, and individual and agency characteristics that influence the dissemination and implementation of EBCDP interventions in Australia, Brazil, China, and the United States.

Methods: Development of the 26-question survey included, a narrative literature review of extant measures in EBCDP; qualitative interviews with 50 chronic disease prevention practitioners in Australia, Brazil, China, and the United States; review by an expert panel of researchers in EBCDP; and test-retest reliability assessment.

Results: A convenience sample of practitioners working in chronic disease prevention in each country completed the survey twice (N = 165). Overall, this tool produced good to moderately reliable responses. Generally, reliability of responses was higher among practitioners from Australia and the United States than China and Brazil.

Conclusions: Reliability findings inform the adaptation and further development of this tool. Revisions to four questions are recommended before use in China and revisions to two questions before use in Brazil. This survey tool can contribute toward an improved understanding of the contextual factors that public health practitioners in Australia,
INTRODUCTION

Chronic diseases are a threat to global health, in developed and developing countries alike, accounting for 60% of deaths worldwide (1). The medical costs and loss of productivity related to chronic diseases are a great financial burden to individuals and economies (1). Evidence-based chronic disease prevention (EBCDP) interventions are effective tools for preventing chronic diseases (2). However, studies among U.S. and European public health practitioners indicate that only 56–64% of chronic disease prevention interventions currently in use are evidence-based (3, 4), while estimates of use of EBCDP interventions in lower and middle income countries are unknown. Studies in Australia and the United States have identified multi-level contextual factors that influence the dissemination and implementation (D&I) of EBCDP interventions. Examples of these contextual factors include individual- and agency-level capacity characterized by the training, structure, material and human resources at hand that hinder or facilitate the use of EBCDP interventions (2, 5–7). Additional work has addressed some of the contextual barriers by training practitioners on the evidence-based decision-making process, specifically clarifying the reasons for selecting EBCDP interventions and outlining how to find the interventions and resources to support effective implementation and quality improvement (3, 4, 7). These studies report increases in the D&I of EBCDP interventions among practitioners who attended the trainings. Research on Canadian public health departments has identified tailored messaging as an effective method for promoting the D&I of evidence-based interventions (8), and examined the pathways through which evidence is shared through organizational systems (9). These contextually specific findings inform next steps in addressing barriers and promoting evidence-based decision-making across the Canada. Little is known about these contextual factors that influence the D&I of EBCDP interventions in developing countries, nor the similarities and differences of contextual factors across countries. Several studies call for global strategies to improve the D&I of EBCDP interventions in order to more effectively reduce chronic diseases around the world (10–12). Reviews of measures used to assess the contextual factors that influence the D&I of EBCDP interventions highlight a lack of psychometric testing of the existing measures and room for improvement among those that have been tested (13–15). To assess cross-country contextual factors and inform globally-focused recommendations for facilitating the D&I of EBCDP interventions, a single survey tool that can be used across multiple, diverse countries is needed.

This study provides a detailed overview of the development and test-retest reliability of a survey tool to measure the stage of dissemination, multi-level contextual factors, and individual and agency characteristics that influence the D&I of EBCDP interventions in Australia, Brazil, China, and the United States. These countries were chosen for several reasons including, their leadership in distinct regions of the world (16–20), differences on contextual variables of interest (e.g., sociocultural, political/economic) (21), and high prevalence of chronic diseases (22). The World Health Organization reports from 2014 showed that the large majority of deaths in each of the four countries was due to chronic diseases (91% in Australia, 88% in the United States, 87% in China, and 74% in Brazil) (22). Further, based on the few studies of the D&I of EBCDP from Brazil and China (23, 24), compared with the many from Australia and the United States (25–29), Brazil, and China were selected as countries likely in earlier stages of dissemination of EBCDP than Australia and the United States.

MATERIALS AND METHODS

Survey Tool Development

Development of the 26-question survey occurred in several stages. First, a guiding framework was developed based on previous work (30, 31) of the research team (see Figure 1). This framework informed subsequent stages of survey tool development, ensuring that qualitative interview questions and initial survey drafts were literature-based and comprehensive from the outset.

Second, a narrative literature review of extant measures in EBCDP was carried out in order to identify relevant questions and gaps in the D&I of EBCDP literature (2, 6, 31–35). Third, between February and July 2015 semi-structured interviews of public health practitioners in Australia (n = 13), Brazil (n = 9), China (n = 16), and the United States (n = 12) were conducted by trained researchers. Practitioners were identified through purposive sampling based on their employment at agencies responsible for the prevention of chronic disease in each country, including community health services, regional health departments, and non-government organizations (Australia); the ministry of health and local health departments (Brazil); hospitals, community health centers, and the Centers for Disease Control and Prevention (China); and local health departments (United States). The interviews were performed in English, Chinese, or Portuguese, audio recorded, transcribed, translated to English by two bi-lingual research team members (n = 25)
Budd et al. Developing Tool to Assess Implementation

FIGURE 1 | Factors affecting the stages of dissemination of evidence-based programs and policies for chronic disease prevention.

when appropriate, and analyzed using deductive, hierarchical coding in NVivo version 10.

Forth, drafts of the survey underwent expert review by 13 chronic disease prevention researchers and were translated forward and backward to Chinese and Portuguese from English. Survey questions were organized into one of the five stages of dissemination or as multi-level contextual factors seen in Figure 1. Individual and agency characteristics were also included. Seven response items were deemed non-applicable or inappropriate for China contexts, but were included in the survey for the other three countries. These response items and the resulting tool can be found in Table 1.

Fifth, research team members in each country recruited public health practitioners working in chronic disease prevention, primarily on the local and regional levels, in each of the four countries to complete the survey. Samples of practitioners from various regions of each country were identified through national databases and networks of chronic disease prevention practitioners between November 2015 and April 2016. Public health systems across countries varied so much that there was no equivalent sampling method that worked for all four countries. In the United States, a stratified (by region) random sample of chronic disease prevention practitioners from a national database received up to three emails and two follow-up telephone calls requesting participation in the electronic survey (58% response rate). In Australia, up to two emails requesting participation in the electronic survey were sent to all chronic disease practitioners in a national registry (18% response rate). In Brazil, the same protocol as was followed in the United States was used, but with an additional follow-up telephone call (46% response rate). In China, a convenience sample of practitioners working within a network of community hospitals received one email and one follow-up telephone call requesting participation in the electronic survey (87% response rate). All surveys were delivered by an email embedded link and completed electronically. Upon completion of the survey, all respondents were asked to retake the survey two to three weeks later for test-retest reliability testing purposes. This process was repeated until each respondent to the survey had been contacted twice, requesting them to retake the survey. Calculating Cohen’s kappa and Intraclass correlation coefficients (ICC) ranging from 0.50 to 0.70 require a sample size of 25–50 test-retest pairs, respectively (38), thus 25 pairs were the minimum, but 50 pairs were the goal. During data collection, political events in Brazil affected the work lives of many Brazilian chronic disease practitioners and made recruitment of Brazilian practitioners extraordinarily difficult (39, 40). The data collection period was extended for research team investigators in Brazil in order to reach the minimum sample size.

This study was carried out in accordance with the committee responsible for human experimentation (institutional and national) and with the World Medical Association’s Declaration of Helsinki with informed consent from all subjects. After reading the electronic informed consent document, subjects indicated their consent by selecting a radial button at the bottom of the informed consent document that read, “I consent to participate in this research study.” Additional written documentation of consent was waived and the protocol was approved by The University of Melbourne Human Ethics Committee, Pontificia Universidade Catolica do Parana Research Ethics Committee, The Hong Kong Polytechnic University Human Ethics Committee of the Faculty of Health and Social
TABLE 1 | Factors influencing the dissemination and implementation of evidence-based chronic disease prevention across four countries: a survey tool.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Response options</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Awareness</strong></td>
<td></td>
</tr>
<tr>
<td>Evidence-based public health is defined as: “the process of integrating science-based interventions with community preferences to improve the health of populations” (36).</td>
<td></td>
</tr>
<tr>
<td>1. With this definition in mind, how knowledgeable are you with evidence-based processes? (select one)</td>
<td></td>
</tr>
<tr>
<td>• Not at all knowledgeable</td>
<td></td>
</tr>
<tr>
<td>• Slightly knowledgeable</td>
<td></td>
</tr>
<tr>
<td>• Somewhat knowledgeable</td>
<td></td>
</tr>
<tr>
<td>• Moderately knowledgeable</td>
<td></td>
</tr>
<tr>
<td>• Extremely knowledgeable</td>
<td></td>
</tr>
<tr>
<td><strong>Adoption</strong></td>
<td></td>
</tr>
<tr>
<td>Definition: Evidence-based interventions are those that several studies have found to be effective at preventing chronic disease. Repositories are collections of evidence-based interventions (e.g., Guide to Community Preventive Services) (US), Health-Evidence.org (Australia), Cochrane Collaboration (US, Australia).</td>
<td></td>
</tr>
<tr>
<td>2. I have used repositories to find evidence-based interventions: (select one)</td>
<td></td>
</tr>
<tr>
<td>• In none of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>• In a few of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>• In many of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>• In all of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>3. Staff at my agency use repositories of evidence-based interventions: (select one)</td>
<td></td>
</tr>
<tr>
<td>• In none of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>• In a few of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>• In many of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>• In all of my programmatic areas</td>
<td></td>
</tr>
<tr>
<td>4. When you make decisions about such things as program planning and implementation, policy development, or funding, which of the following are important to you? (select the top three)</td>
<td></td>
</tr>
<tr>
<td>• Support from leadership at my agency</td>
<td></td>
</tr>
<tr>
<td>• Support from elected officials</td>
<td></td>
</tr>
<tr>
<td>• Support from community partnerships</td>
<td></td>
</tr>
<tr>
<td>• Recommendations from the funding agency</td>
<td></td>
</tr>
<tr>
<td>• Colleagues are using the intervention</td>
<td></td>
</tr>
<tr>
<td>• Available resources (program dollars and staff)</td>
<td></td>
</tr>
<tr>
<td>• How easy the intervention or policy is to implement</td>
<td></td>
</tr>
<tr>
<td>• Evidence regarding the effectiveness of the intervention</td>
<td></td>
</tr>
<tr>
<td>• Health planning tools (e.g., MAPP or Health People 2010)</td>
<td></td>
</tr>
<tr>
<td>• Relevance of the intervention to the population of interest</td>
<td></td>
</tr>
<tr>
<td>• Seriousness of the health problem</td>
<td></td>
</tr>
<tr>
<td>• Other, please specify _____</td>
<td></td>
</tr>
<tr>
<td>• Not applicable</td>
<td></td>
</tr>
<tr>
<td>5. What avenues do you use to learn about the current study findings on evidence-based chronic disease prevention interventions? (select all that apply)</td>
<td></td>
</tr>
<tr>
<td>• Academic journals</td>
<td></td>
</tr>
<tr>
<td>• Conferences</td>
<td></td>
</tr>
<tr>
<td>• Email alerts</td>
<td></td>
</tr>
<tr>
<td>• Evidence-based repositories</td>
<td></td>
</tr>
<tr>
<td>• Facebook</td>
<td></td>
</tr>
<tr>
<td>• Funders⁹</td>
<td></td>
</tr>
<tr>
<td>• Government agency staff</td>
<td></td>
</tr>
<tr>
<td>• Government reports</td>
<td></td>
</tr>
<tr>
<td>• Internet search engines</td>
<td></td>
</tr>
<tr>
<td>• Listservs/Newsletters/Online forums</td>
<td></td>
</tr>
<tr>
<td>• Media campaigns/Media interviews</td>
<td></td>
</tr>
<tr>
<td>• Networks</td>
<td></td>
</tr>
<tr>
<td>• Partnerships (e.g., with universities, health departments, professional associations)</td>
<td></td>
</tr>
<tr>
<td>• Policy briefs⁹</td>
<td></td>
</tr>
<tr>
<td>• Press releases</td>
<td></td>
</tr>
<tr>
<td>• Stakeholders⁹</td>
<td></td>
</tr>
<tr>
<td>• Technical assistance/Data liaison</td>
<td></td>
</tr>
<tr>
<td>• Trainings/Workshops/Meetings within my agency</td>
<td></td>
</tr>
<tr>
<td>• Webinars</td>
<td></td>
</tr>
<tr>
<td>• Other, please specify _____</td>
<td></td>
</tr>
<tr>
<td>• None</td>
<td></td>
</tr>
<tr>
<td>6. For which avenues would you like additional access? (select all that apply)</td>
<td></td>
</tr>
<tr>
<td>Same responses as #13</td>
<td></td>
</tr>
<tr>
<td><strong>Implementation</strong></td>
<td></td>
</tr>
<tr>
<td>7. Approximately what percentage of programs supported by your agency would you say are evidence-based?</td>
<td></td>
</tr>
<tr>
<td>Fill in the blank 0–100%</td>
<td></td>
</tr>
<tr>
<td>8. As you think about the future, what is one thing you would change to help you implement evidence-based chronic disease prevention interventions?</td>
<td></td>
</tr>
<tr>
<td>Fill in the blank</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
### TABLE 1 | Continued

<table>
<thead>
<tr>
<th>Questions</th>
<th>Response options</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maintenance</strong></td>
<td></td>
</tr>
<tr>
<td>Quality improvement (QI) refers to ongoing formal assessments of the effectiveness and quality of public health chronic disease prevention efforts. (37). Some examples of quality improvement processes include: Results-based accountability (RBA), Community Health Improvement Plan (CHIP), Plan-Do-Study-Act (PDSA), and Plan-Do-Check-Act.</td>
<td>• In none of my programmatic areas</td>
</tr>
<tr>
<td>9. Staff at my agency use quality improvement processes: (select one)</td>
<td>• In a few of my programmatic areas</td>
</tr>
<tr>
<td>10. In your opinion, how often do programs end that should have continued? (i.e., end without warrant) (select one)</td>
<td>• In many of my programmatic areas</td>
</tr>
<tr>
<td>11. When you think about public health programs that have ended, what are the most common reasons for programs ending? (Select the top three)</td>
<td>• In all of my programmatic areas</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Contextual factors</strong></td>
<td></td>
</tr>
<tr>
<td>14. Which of the following are personal barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions? (Select all that apply)</td>
<td>• Never</td>
</tr>
<tr>
<td></td>
<td>• Sometimes</td>
</tr>
<tr>
<td></td>
<td>• Often</td>
</tr>
<tr>
<td>15. Which of the following are agency-level barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions? (Select all that apply)</td>
<td>• Program was never evaluated</td>
</tr>
<tr>
<td></td>
<td>• Program was evaluated but did not demonstrate impact</td>
</tr>
<tr>
<td></td>
<td>• Opposition/lack of support from leaders in my agency</td>
</tr>
<tr>
<td></td>
<td>• Opposition/lack of support from the general public</td>
</tr>
<tr>
<td></td>
<td>• Opposition/lack of support from policy makers</td>
</tr>
<tr>
<td></td>
<td>• Funding diverted to a higher priority program</td>
</tr>
<tr>
<td></td>
<td>• Grant funding ended</td>
</tr>
<tr>
<td></td>
<td>• Change in political leadership</td>
</tr>
<tr>
<td></td>
<td>• Insurance funding/coverage ended</td>
</tr>
<tr>
<td></td>
<td>• Program was adopted or continued by other organizations</td>
</tr>
<tr>
<td></td>
<td>• A program champion departed</td>
</tr>
<tr>
<td></td>
<td>• Program was not evidence-based</td>
</tr>
<tr>
<td></td>
<td>• Program was expensive</td>
</tr>
<tr>
<td></td>
<td>• Program was challenging to maintain</td>
</tr>
<tr>
<td></td>
<td>• Other, please specify _____</td>
</tr>
<tr>
<td></td>
<td>• I do not know</td>
</tr>
<tr>
<td></td>
<td>• Not applicable</td>
</tr>
<tr>
<td>12. In your opinion, how often do programs continue that should have ended? (i.e., continue without warrant) (select one)</td>
<td>• Never</td>
</tr>
<tr>
<td></td>
<td>• Sometimes</td>
</tr>
<tr>
<td></td>
<td>• Often</td>
</tr>
<tr>
<td>13. When you think about public health programs that continued that should have ended, what are the most common reasons for their continuation? (i.e., continue without warrant) (Select the top three)</td>
<td>• Program was never evaluated</td>
</tr>
<tr>
<td></td>
<td>• Sustained support from leaders in your agency</td>
</tr>
<tr>
<td></td>
<td>• Sustained support from the general public</td>
</tr>
<tr>
<td></td>
<td>• Sustained support from policymakers</td>
</tr>
<tr>
<td></td>
<td>• Prohibitive costs of starting something new</td>
</tr>
<tr>
<td></td>
<td>• Absence of alternative options</td>
</tr>
<tr>
<td></td>
<td>• Sustained funding</td>
</tr>
<tr>
<td></td>
<td>• Presence of a program champion</td>
</tr>
<tr>
<td></td>
<td>• Program was considered evidence-based</td>
</tr>
<tr>
<td></td>
<td>• Program was low-cost</td>
</tr>
<tr>
<td></td>
<td>• Program was easy to maintain</td>
</tr>
<tr>
<td></td>
<td>• Other, please specify _____</td>
</tr>
<tr>
<td></td>
<td>• I do not know</td>
</tr>
<tr>
<td></td>
<td>• Not applicable</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Continued</strong></td>
<td></td>
</tr>
<tr>
<td>14. Which of the following are personal barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions? (Select all that apply)</td>
<td>• Not being an expert on relevant issues</td>
</tr>
<tr>
<td></td>
<td>• Lack of confidence in finding data and statistics</td>
</tr>
<tr>
<td></td>
<td>• Lack of skills to develop evidence-based interventions</td>
</tr>
<tr>
<td></td>
<td>• Lack of confidence in carrying out evidence-based interventions</td>
</tr>
<tr>
<td></td>
<td>• Lack of decision-making authority</td>
</tr>
<tr>
<td></td>
<td>• Low value of evidence-based approaches</td>
</tr>
<tr>
<td></td>
<td>• Workload is too heavy/not enough time</td>
</tr>
<tr>
<td></td>
<td>• Overwhelmed by task</td>
</tr>
<tr>
<td></td>
<td>• Other, please specify _____</td>
</tr>
<tr>
<td></td>
<td>• None</td>
</tr>
<tr>
<td>15. Which of the following are agency-level barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions? (Select all that apply)</td>
<td>• Poor understanding of evidence-based approaches</td>
</tr>
<tr>
<td></td>
<td>• Culture/climate is not supportive of change/new ideas</td>
</tr>
<tr>
<td></td>
<td>• No existing policies to support evidence-based approaches</td>
</tr>
<tr>
<td></td>
<td>• Agency does not provide training in evidence-based approaches</td>
</tr>
</tbody>
</table>

(Continued)
TABLE 1 | Continued

<table>
<thead>
<tr>
<th>Questions</th>
<th>Response options</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Staff/leaders lack formal training in evidence-based approaches</td>
<td>• Lack of access to resources (e.g., computer, Internet)</td>
</tr>
<tr>
<td>• Low priority placed on chronic disease prevention</td>
<td>• Staff/leaders lack formal training in evidence-based approaches</td>
</tr>
<tr>
<td>• Not enough funding</td>
<td>• Lack of access to resources (e.g., computer, Internet)</td>
</tr>
<tr>
<td>• No systems to ensure interventions are evidence-based</td>
<td>• Not enough funding</td>
</tr>
<tr>
<td>• Not enough staff</td>
<td>• Low priority placed on chronic disease prevention</td>
</tr>
<tr>
<td>• Beliefs that evidence-based interventions are too difficult to</td>
<td>• No systems to ensure interventions are evidence-based</td>
</tr>
<tr>
<td>implement/sustain</td>
<td>• Not enough staff</td>
</tr>
<tr>
<td>• Other, please specify ______</td>
<td>• Low priority placed on chronic disease prevention</td>
</tr>
<tr>
<td>• None</td>
<td>• Not enough staff</td>
</tr>
<tr>
<td>• Lack of access to repositories/databases of scientific studies</td>
<td>• Beliefs that evidence-based interventions are too difficult to implement/sustain</td>
</tr>
<tr>
<td>• Lack of partnership between agency and community</td>
<td>• Other, please specify ______</td>
</tr>
<tr>
<td>• Community members’ needs compete with evidence-based recommendations</td>
<td>• None</td>
</tr>
<tr>
<td>• Catering to preferences of funders⁹</td>
<td>• None</td>
</tr>
<tr>
<td>• Low priority placed on chronic disease prevention</td>
<td>• Other, please specify ______</td>
</tr>
<tr>
<td>• Not enough staff</td>
<td>• None</td>
</tr>
<tr>
<td>• Evidence is presented in a language I do not understand</td>
<td>• None</td>
</tr>
<tr>
<td>• Community cultural practices conflict with evidence-based recommendations</td>
<td>• None</td>
</tr>
<tr>
<td>• Not enough relevant evidence for populations served</td>
<td>• Distrust of scientific data in the populations served</td>
</tr>
<tr>
<td>• Serving a rural setting where data are lacking⁸</td>
<td>• Community cultural practices conflict with evidence-based recommendations</td>
</tr>
<tr>
<td>• Serving a highly disadvantaged population</td>
<td>• Not enough relevant evidence for populations served</td>
</tr>
<tr>
<td>• Serving a population that speaks a language different from the majority⁸</td>
<td>• Serving a highly disadvantaged population</td>
</tr>
<tr>
<td>• Evidence is presented in a language I do not understand</td>
<td>• Serving a population that speaks a language different from the majority⁸</td>
</tr>
<tr>
<td>• Other, please specify ______</td>
<td>• Evidence is presented in a language I do not understand</td>
</tr>
<tr>
<td>• None</td>
<td>• Other, please specify ______</td>
</tr>
<tr>
<td>• Political leaders not providing enough support</td>
<td>• None</td>
</tr>
<tr>
<td>• Funding changes that occur with changes in political leadership</td>
<td>• Political leaders not providing enough support</td>
</tr>
<tr>
<td>• Political climate conflicts with evidence-based chronic disease prevention recommendations</td>
<td>• Funding changes that occur with changes in political leadership</td>
</tr>
<tr>
<td>• Health care system does not support evidence-based chronic disease prevention</td>
<td>• Political climate conflicts with evidence-based chronic disease prevention</td>
</tr>
<tr>
<td>• Other, please specify ______</td>
<td>• Health care system does not support evidence-based chronic disease prevention</td>
</tr>
<tr>
<td>• None</td>
<td>• Other, please specify ______</td>
</tr>
<tr>
<td>• Prioritizing program and policy options</td>
<td>• None</td>
</tr>
<tr>
<td>• Quantifying the public health issue using descriptive epidemiology (e.g., concepts of person, place, time)</td>
<td>• Prioritizing program and policy options</td>
</tr>
<tr>
<td>• Using quantitative evaluation approaches (e.g., surveillance or surveys)</td>
<td>• Using quantitative evaluation approaches (e.g., surveillance or surveys)</td>
</tr>
<tr>
<td>• Using qualitative evaluation approaches (e.g., focus groups, key informant interviews)</td>
<td>• Using quantitative evaluation approaches (e.g., surveillance or surveys)</td>
</tr>
<tr>
<td>• Developing an action plan for achieving goals</td>
<td>• Using qualitative evaluation approaches (e.g., focus groups, key informant interviews)</td>
</tr>
<tr>
<td>• Defining the health issue according to the community's needs and assets</td>
<td>• Developing an action plan for achieving goals</td>
</tr>
<tr>
<td>• Adapting interventions for different communities and settings</td>
<td>• Defining the health issue according to the community's needs and assets</td>
</tr>
<tr>
<td>• Using economic data in the decision making process</td>
<td>• Adapting interventions for different communities and settings</td>
</tr>
<tr>
<td>• Communicating research to policy makers</td>
<td>• Using economic data in the decision making process</td>
</tr>
<tr>
<td>• Other, please specify ______</td>
<td>• Communicating research to policy makers</td>
</tr>
<tr>
<td>• None</td>
<td>• Other, please specify ______</td>
</tr>
</tbody>
</table>

16. Which of the following are community-level barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions? (Select all that apply)

17. Which of the following are sociocultural barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions? (Select all that apply)

18. Which of the following are political/economic barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions? (Select all that apply)

19. For which of the following skills would you like additional technical support or training? (Check all that apply)

Individual and agency characteristics

20. What is your gender? (select one)

21. What is your age? (select one)
<table>
<thead>
<tr>
<th>Questions</th>
<th>Response options</th>
</tr>
</thead>
<tbody>
<tr>
<td>22. What degree/credentials do you hold? (Check all that apply)</td>
<td>• BS/BA</td>
</tr>
<tr>
<td></td>
<td>• CHES</td>
</tr>
<tr>
<td></td>
<td>• Certified Health Educator (in Diabetes, Asthma, etc.)</td>
</tr>
<tr>
<td></td>
<td>• RN or RD</td>
</tr>
<tr>
<td></td>
<td>• MS or MSc</td>
</tr>
<tr>
<td></td>
<td>• MPH or MSPH</td>
</tr>
<tr>
<td></td>
<td>• MA</td>
</tr>
<tr>
<td></td>
<td>• Other Master’s degree</td>
</tr>
<tr>
<td></td>
<td>• NP</td>
</tr>
<tr>
<td></td>
<td>• MO or DO</td>
</tr>
<tr>
<td></td>
<td>• PhD, DrPH, ScD</td>
</tr>
<tr>
<td></td>
<td>• Other, please specify ______</td>
</tr>
<tr>
<td>23. Though you may work in several capacities, how do you best describe your primary position? (select one)</td>
<td>• Academic Researcher</td>
</tr>
<tr>
<td></td>
<td>• Academic Educator</td>
</tr>
<tr>
<td></td>
<td>• Community Health Nurse</td>
</tr>
<tr>
<td></td>
<td>• Department Head</td>
</tr>
<tr>
<td></td>
<td>• Division or Bureau Head/ Division Deputy</td>
</tr>
<tr>
<td></td>
<td>• Director</td>
</tr>
<tr>
<td></td>
<td>• Epidemiologist</td>
</tr>
<tr>
<td></td>
<td>• Health Educator</td>
</tr>
<tr>
<td></td>
<td>• Nutritionist/Dietician</td>
</tr>
<tr>
<td></td>
<td>• Physician</td>
</tr>
<tr>
<td></td>
<td>• Program Manager/Administrator/Coordinator</td>
</tr>
<tr>
<td></td>
<td>• Program Planner/ Evaluator</td>
</tr>
<tr>
<td></td>
<td>• Public Health Specialist</td>
</tr>
<tr>
<td></td>
<td>• Social Worker</td>
</tr>
<tr>
<td></td>
<td>• Statisticist</td>
</tr>
<tr>
<td></td>
<td>• Other, please specify ______</td>
</tr>
<tr>
<td>24. The agency in which I work has the following number of employees. (select one)</td>
<td>• 0–50</td>
</tr>
<tr>
<td></td>
<td>• 51–100</td>
</tr>
<tr>
<td></td>
<td>• 101–200</td>
</tr>
<tr>
<td></td>
<td>• 201–400</td>
</tr>
<tr>
<td></td>
<td>• 401–800</td>
</tr>
<tr>
<td></td>
<td>• &gt;800</td>
</tr>
<tr>
<td></td>
<td>• I do not know</td>
</tr>
<tr>
<td>25. The size of the population my agency serves is has the following number of people. (select one)</td>
<td>• 0–24,999</td>
</tr>
<tr>
<td></td>
<td>• 25,000–49,999</td>
</tr>
<tr>
<td></td>
<td>• 50,000–74,999</td>
</tr>
<tr>
<td></td>
<td>• 75,000–99,999</td>
</tr>
<tr>
<td></td>
<td>• 100,000–149,999</td>
</tr>
<tr>
<td></td>
<td>• 150,000–199,999</td>
</tr>
<tr>
<td></td>
<td>• 200,000–299,999</td>
</tr>
<tr>
<td></td>
<td>• 300,000–399,999</td>
</tr>
<tr>
<td></td>
<td>• 400,000+</td>
</tr>
<tr>
<td></td>
<td>• I do not know</td>
</tr>
<tr>
<td>26. Is there anything else you would like to share on the topic of evidence-based chronic disease prevention? Please specify</td>
<td>Fill in the blank</td>
</tr>
</tbody>
</table>

*This item was not applicable and not included in the survey for respondents in China.*

Science, and Washington University in St. Louis Institutional Review Board.

**Analyses**

Test-retest reliability was examined on the survey questions, excluding open-ended questions and individual and agency characteristics. Intraclass correlation coefficients (ICC) were calculated for questions with ordinal response options (questions 1 through 3, 9, 10, and 12; see Table 1). “I don’t know” and “not applicable” response options were not included in the ICC calculations. Each response item for questions 4, 5, 11, and 13 through 19 was dichotomized to reflect whether a respondent selected the response option or not. Cohen’s kappa was run for each of these response options individually. The mean of all of the Cohen’s kappas for each question’s set of response options was calculated. Cut-points for ICC and mean kappa (excellent: ≥0.801; good: 0.601–0.80; moderate: 0.401–0.60; poor: ≤0.40) were selected based on recommendations (41, 42), and to aid in the interpretation of the results. Percentage agreement was also calculated for all of the aforementioned questions, excluding question 7, which asked respondents to provide a percentage. Questions for which...
mean kappa was calculated, mean percentage agreement was also calculated. Cut-points for percentage agreement included: excellent: 89.5–100%; good: 74.5–89.4%; moderate: 60–74.4%; and poor: <60%. All analyses were conducted in Stata version 14.

RESULTS

There were 400 survey respondents total and 165 of them took the survey twice for test-retest reliability purposes (N = 39 from Australia; N = 27 from Brazil; N = 45 from China; N = 54 from the United States). The test-retest respondents were all public health practitioners (e.g., nutritionist/dietician, coordinator, community health nurse) working in chronic disease prevention. Public Health Specialist was added as a primary employment position option post hoc, in order to capture a common “other” response provided by practitioners from Brazil. Respondents were primarily female (79%) between 30 and 49 years old (53%). The mean survey completion time varied by country, with Brazil having the longest (33.2 min ± 27.8), followed by the United States (17.72 min ± 13.4), Australia (16.6 min ± 10.0), and China (13.8 min ± 10.5). The mean number of days between test and retest was greatest in Brazil (46.4 ± 28.5), followed by Australia (39.0 ± 2.8), China (23.7 ± 7.6) and the United States (21.0 ± 9.1). Table 2 shows frequency counts for each response option by country, the first time respondents completed the survey. Item responses vary in prevalence from zero endorsements to endorsement from a large majority of a country’s sample.

The test-retest reliability coefficients and percentage agreement by question and country appear in Table 3. Of the seven questions with ordinal response options assessed using ICC, six and seven demonstrated good to moderate reliability among practitioners from Australia and the United States, respectively, whereas three questions among practitioners from Brazil and China demonstrated good to moderate reliability. Six of those seven questions were also assessed using percentage agreement. Six and five of the questions demonstrated good to moderate percentage agreement among practitioners from Australia and the United States, respectively, whereas three questions among practitioners from Brazil and one among practitioners from China demonstrated moderate percentage agreement at best.

Of the 11 questions whose response options were dichotomized and assessed using mean Cohen’s kappa, few questions among practitioners across all four countries showed moderate mean reliability at best (Australia, N = 2; Brazil, N = 1; China, N = 1; United States, N = 3). Mean percentage agreement told a different story for these 11 questions. All but one question showed good mean percentage agreement among practitioners from Australia and the United States. Seven and five questions showed good mean percentage agreement among practitioners from Brazil and China, respectively. The remaining of the 11 questions across the countries showed moderate mean percentage agreement.

The following four questions produced less than moderately reliable responses based on both ICC and percentage agreement among practitioners in China: Personal use of repositories to find evidence-based interventions; Workplace staff use of repositories to find evidence-based interventions; Frequency that programs end that should have continued; and Frequency that programs continue that should have ended. Two of those questions (Workplace staff use of repositories to find evidence-based interventions, and Frequency that programs end that should have continued) produced less than moderately reliable responses among practitioners from Brazil based on both measures of reliability as well.

DISCUSSION

The development and reliability testing of this survey tool are important early steps toward facilitating population-level research that can increase our knowledge of country-specific and cross-country contextual factors that influence the D&I of EBCDP interventions and, in turn, begin to inform more global strategies for improving the D&I of EBCDP. This study, novel in its common methods across countries, showed that the measurement tool produced moderate to good reliability of responses, with at least one measure of reliability, among 14 of the 18 questions across all four countries.

Reliability findings inform the adaptation and further development of this tool. For example, the authors recommend revising the four questions pertaining to personal and workplace staff use of repositories for finding evidence-based interventions and frequency that programs end or continue without warrant before further use among practitioners in China and Brazil. The poor reliability of responses produced from these questions among practitioners from Brazil and China reflect a difference in how they relate to the content of the questions, compared with practitioners from Australia and the United States. This difference may highlight meaningful differences within contexts with respect to D&I processes and structures. For instance, practitioners in countries for which EBCDP is in an earlier stage of dissemination tend to be less knowledgeable about key concepts of EBCDP, making the questions conceptually more difficult and in turn negatively influencing the reliability of their responses (43). Another potential contributing factor to the lower reliability among responses from practitioners in Brazil and China is that the survey tool had to be translated from English to Chinese and Portuguese. Tanzer and Sim review international guidelines on translating and adapting measures across cultural contexts, and this study reflects well the best practices for developing a relevant survey tool for use in the four intended countries (44). For instance, bilingual researchers from each of the four cultural perspectives, as well as public health practitioners working in the chronic disease prevention context in each country were involved in the development of the questions, response options, translations, and reliability testing. Despite steps that the research team took to minimize mis-translation, the meaning of each question and response option becomes one layer removed from its original, intended meaning after translation. Next steps for informing further adaptation of the survey tool should include validity testing.
**TABLE 2 | Frequency of response option endorsement by country (N = 165).**

<table>
<thead>
<tr>
<th>Question and response options</th>
<th>Australia (Total N = 39)</th>
<th>Brazil (Total N = 27)</th>
<th>China (Total N = 45)</th>
<th>United States (Total N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>1. How knowledgeable are you with evidence-based processes?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at all knowledgeable</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>7 (15.6)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Slightly knowledgeable</td>
<td>0 (0.0)</td>
<td>1 (3.7)</td>
<td>15 (33.3)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Somewhat knowledgeable</td>
<td>9 (23.1)</td>
<td>8 (29.6)</td>
<td>16 (35.6)</td>
<td>6 (11.1)</td>
</tr>
<tr>
<td>Moderately knowledgeable</td>
<td>22 (56.4)</td>
<td>13 (48.1)</td>
<td>8 (17.8)</td>
<td>31 (57.4)</td>
</tr>
<tr>
<td>Extremely knowledgeable</td>
<td>8 (20.5)</td>
<td>5 (18.5)</td>
<td>1 (2.2)</td>
<td>16 (29.6)</td>
</tr>
<tr>
<td>2. I have used repositories to find evidence-based interventions:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In none of my programmatic areas</td>
<td>2 (5.1)</td>
<td>0 (0.0)</td>
<td>12 (26.7)</td>
<td>3 (5.6)</td>
</tr>
<tr>
<td>In a few of my programmatic areas</td>
<td>11 (28.2)</td>
<td>7 (25.9)</td>
<td>25 (55.6)</td>
<td>18 (33.3)</td>
</tr>
<tr>
<td>In many of my programmatic areas</td>
<td>16 (41.0)</td>
<td>7 (25.9)</td>
<td>6 (13.3)</td>
<td>27 (50.0)</td>
</tr>
<tr>
<td>In all of my programmatic areas</td>
<td>8 (20.5)</td>
<td>13 (48.1)</td>
<td>1 (2.2)</td>
<td>5 (9.3)</td>
</tr>
<tr>
<td>I don’t know</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Not applicable</td>
<td>2 (5.13)</td>
<td>0 (0.0)</td>
<td>3 (6.7)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>3. Staff at my agency use repositories of evidence-based interventions:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In none of their programmatic areas</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>11 (24.4)</td>
<td>2 (3.7)</td>
</tr>
<tr>
<td>In a few of their programmatic areas</td>
<td>9 (23.1)</td>
<td>5 (18.5)</td>
<td>19 (42.2)</td>
<td>18 (33.3)</td>
</tr>
<tr>
<td>In many of their programmatic areas</td>
<td>19 (48.7)</td>
<td>11 (40.7)</td>
<td>8 (17.8)</td>
<td>24 (44.4)</td>
</tr>
<tr>
<td>In all of their programmatic areas</td>
<td>4 (10.3)</td>
<td>9 (33.3)</td>
<td>0 (0.0)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>I don’t know</td>
<td>3 (7.7)</td>
<td>1 (3.7)</td>
<td>9 (20.0)</td>
<td>7 (13.0)</td>
</tr>
<tr>
<td>Not applicable</td>
<td>3 (7.7)</td>
<td>1 (3.7)</td>
<td>0 (0.0)</td>
<td>2 (3.7)</td>
</tr>
<tr>
<td>4. When you make decisions about such things as program planning and implementation, policy development, or funding, which of the following are important to you?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support from leadership at my agency</td>
<td>9 (23.1)</td>
<td>22 (81.5)</td>
<td>31 (68.9)</td>
<td>24 (44.4)</td>
</tr>
<tr>
<td>Support from elected officials</td>
<td>5 (12.8)</td>
<td>13 (48.1)</td>
<td>17 (37.8)</td>
<td>4 (7.4)</td>
</tr>
<tr>
<td>Support from community partnerships</td>
<td>13 (33.3)</td>
<td>20 (74.1)</td>
<td>16 (35.6)</td>
<td>21 (38.9)</td>
</tr>
<tr>
<td>Recommendations from the funding agency/ Recommendations from the Research Management Department (China)</td>
<td>1 (2.6)</td>
<td>17 (63.0)</td>
<td>10 (22.2)</td>
<td>16 (29.6)</td>
</tr>
<tr>
<td>Colleagues are using the intervention</td>
<td>1 (2.6)</td>
<td>19 (70.4)</td>
<td>6 (13.3)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Available resources (program dollars and staff)</td>
<td>15 (38.5)</td>
<td>26 (96.3)</td>
<td>25 (55.6)</td>
<td>36 (66.7)</td>
</tr>
<tr>
<td>How easy the intervention or policy is to implement</td>
<td>2 (5.1)</td>
<td>12 (44.4)</td>
<td>23 (51.1)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Evidence regarding the effectiveness of the intervention</td>
<td>30 (76.9)</td>
<td>23 (85.2)</td>
<td>13 (28.9)</td>
<td>32 (59.3)</td>
</tr>
<tr>
<td>Health planning tools (e.g. MAPP or Health People 2010)/Government Health plans (China)</td>
<td>5 (12.8)</td>
<td>21 (77.8)</td>
<td>9 (20.0)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Relevance of the intervention to the population of interest</td>
<td>26 (66.7)</td>
<td>24 (88.9)</td>
<td>18 (40.0)</td>
<td>20 (37.0)</td>
</tr>
<tr>
<td>Seriousness of the health problem</td>
<td>9 (23.1)</td>
<td>20 (74.1)</td>
<td>9 (20.0)</td>
<td>5 (9.3)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (2.6)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Not applicable</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1 (2.2)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>5. What avenues do you use to learn about the current study findings on evidence-based chronic disease prevention interventions?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic journals</td>
<td>36 (92.3)</td>
<td>19 (70.4)</td>
<td>23 (51.1)</td>
<td>24 (44.4)</td>
</tr>
<tr>
<td>Conferences</td>
<td>35 (89.7)</td>
<td>14 (51.9)</td>
<td>14 (31.1)</td>
<td>36 (66.7)</td>
</tr>
<tr>
<td>Email alerts</td>
<td>28 (71.8)</td>
<td>5 (18.5)</td>
<td>5 (11.1)</td>
<td>32 (59.3)</td>
</tr>
<tr>
<td>Evidence-based repositories</td>
<td>17 (43.6)</td>
<td>19 (70.4)</td>
<td>8 (17.8)</td>
<td>31 (57.4)</td>
</tr>
<tr>
<td>Facebook/Weibo, Wechat (China)</td>
<td>4 (10.3)</td>
<td>1 (3.7)</td>
<td>10 (22.2)</td>
<td>4 (7.4)</td>
</tr>
<tr>
<td>Funders³</td>
<td>4 (10.3)</td>
<td>0 (0.0)</td>
<td>–</td>
<td>23 (42.6)</td>
</tr>
<tr>
<td>Government agency staff</td>
<td>12 (30.8)</td>
<td>17 (63.0)</td>
<td>2 (4.4)</td>
<td>25 (46.3)</td>
</tr>
</tbody>
</table>

(Continued)
TABLE 2 | Continued

<table>
<thead>
<tr>
<th>Question and response options</th>
<th>Australia (Total N = 39)</th>
<th>Brazil (Total N = 27)</th>
<th>China (Total N = 45)</th>
<th>United States (Total N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>Government reports</td>
<td>23 (59.0)</td>
<td>23 (85.2)</td>
<td>3 (6.7)</td>
<td>23 (42.6)</td>
</tr>
<tr>
<td>Internet search engines</td>
<td>21 (53.8)</td>
<td>16 (59.3)</td>
<td>14 (31.1)</td>
<td>32 (59.3)</td>
</tr>
<tr>
<td>Listservs/Newsletters/Online forums</td>
<td>15 (38.5)</td>
<td>2 (7.4)</td>
<td>3 (6.7)</td>
<td>25 (46.3)</td>
</tr>
<tr>
<td>Media campaigns/Media interviews</td>
<td>4 (10.3)</td>
<td>3 (11.1)</td>
<td>6 (13.3)</td>
<td>7 (13.0)</td>
</tr>
<tr>
<td>Networks</td>
<td>23 (59.0)</td>
<td>10 (37.0)</td>
<td>5 (11.1)</td>
<td>18 (33.3)</td>
</tr>
<tr>
<td>Partnerships (e.g., with universities, health departments, professional associations)</td>
<td>26 (66.7)</td>
<td>14 (51.9)</td>
<td>7 (15.6)</td>
<td>35 (64.8)</td>
</tr>
<tr>
<td>Policy briefs&lt;sup&gt;a&lt;/sup&gt;</td>
<td>11 (28.2)</td>
<td>12 (44.4)</td>
<td>–</td>
<td>17 (31.5)</td>
</tr>
<tr>
<td>Press releases</td>
<td>8 (20.5)</td>
<td>3 (11.1)</td>
<td>6 (13.3)</td>
<td>9 (16.7)</td>
</tr>
<tr>
<td>Stakeholders&lt;sup&gt;a&lt;/sup&gt;</td>
<td>11 (28.2)</td>
<td>27 (100.0)</td>
<td>–</td>
<td>14 (25.9)</td>
</tr>
<tr>
<td>Technical assistance/Data liaison</td>
<td>1 (2.6)</td>
<td>13 (48.1)</td>
<td>2 (4.4)</td>
<td>21 (38.9)</td>
</tr>
<tr>
<td>Trainings/Workshops/Meetings within my agency</td>
<td>17 (43.6)</td>
<td>11 (40.7)</td>
<td>18 (40.0)</td>
<td>7 (13.0)</td>
</tr>
<tr>
<td>Webinars</td>
<td>16 (41.0)</td>
<td>2 (7.4)</td>
<td>1 (2.2)</td>
<td>36 (66.7)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0.0)</td>
<td>2 (7.4)</td>
<td>0 (0.0)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>None</td>
<td>1 (2.6)</td>
<td>1 (3.7)</td>
<td>5 (11.1)</td>
<td>0 (0.0)</td>
</tr>
</tbody>
</table>

6. For which avenues would you like additional access?

| Academic journals            | 12 (30.8)               | 16 (59.3)             | 17 (37.8)            | 12 (22.2)                   |
| Conferences                  | 8 (20.5)                | 17 (63.0)             | 10 (22.2)            | 16 (29.6)                   |
| Email alerts                 | 3 (7.7)                 | 5 (18.5)              | 10 (22.2)            | 4 (7.4)                     |
| Evidence-based repositories  | 13 (33.3)               | 7 (25.9)              | 19 (42.2)            | 18 (33.3)                   |
| Facebook/Weibo, Wechat (China) | 2 (5.1)              | 2 (7.4)               | 12 (26.7)            | 3 (5.6)                     |
| Funders<sup>a</sup>          | 4 (10.3)                | 3 (11.1)              | –                    | 10 (18.5)                   |
| Government agency staff      | 6 (15.4)                | 2 (7.4)               | 2 (4.4)              | 6 (11.1)                    |
| Government reports           | 8 (20.5)                | 5 (18.5)              | 8 (17.8)             | 3 (5.6)                     |
| Internet search engines      | 1 (2.6)                 | 2 (7.4)               | 11 (24.4)            | 3 (5.6)                     |
| Listservs/Newsletters/Online forums | 4 (10.3)           | 3 (11.1)              | 14 (31.1)            | 8 (14.8)                    |
| Media campaigns/Media interviews | 2 (5.1)             | 1 (3.7)               | 4 (8.9)              | 4 (7.4)                     |
| Networks                     | 9 (23.1)                | 6 (22.2)              | 2 (4.4)              | 8 (14.8)                    |
| Partnerships (e.g., with universities, health departments, professional associations) | 13 (33.3) | 11 (40.7) | 13 (28.9) | 15 (27.8) |
| Policy briefs<sup>a</sup>    | 5 (12.8)                | 2 (7.4)               | 0 (0.0)              | 4 (7.4)                     |
| Press releases               | 2 (5.1)                 | 2 (7.4)               | 7 (15.6)             | 2 (3.7)                     |
| Stakeholders<sup>a</sup>     | 4 (10.3)                | 2 (7.4)               | –                    | 3 (5.6)                     |
| Technical assistance/Data liaison | 6 (15.4)            | 1 (3.7)               | 9 (20.0)             | 12 (22.2)                   |
| Trainings/Workshops/Meetings within my agency | 8 (20.5) | 10 (37.0) | 18 (40.0) | 12 (22.2) |
| Webinars                     | 10 (25.6)               | 1 (3.7)               | 3 (6.7)              | 9 (16.7)                    |
| Other                        | 3 (7.7)                 | 0 (0.0)               | 0 (0.0)              | 0 (0.0)                     |
| None                         | 1 (2.6)                 | 0 (0.0)               | 3 (6.7)              | 0 (0.0)                     |

9. Staff at my agency use quality improvement processes:

| In none of their programmatic areas | 0 (0.0) | 1 (3.7) | 9 (20.0) | 0 (0.0) |
| In a few of their programmatic areas | 10 (25.6) | 8 (29.6) | 16 (35.6) | 20 (37.0) |
| In many of their programmatic areas | 16 (41.0) | 12 (44.4) | 15 (33.3) | 28 (51.9) |
| In all of their programmatic areas | 8 (20.5) | 5 (18.5) | 1 (2.2) | 4 (7.4) |
| I don’t know                      | 3 (7.7) | 1 (3.7) | 5 (11.1) | 1 (1.9) |
| Not applicable                    | 2 (5.1) | 0 (0.0) | 0 (0.0) | 1 (1.9) |

10. In your opinion, how often do programs end that should have continued?

| Never                           | 2 (5.1) | 21 (77.8) | 21 (46.7) | 0 (0.0) |

(Continued)
TABLE 2 | Continued

<table>
<thead>
<tr>
<th>Question and response options</th>
<th>Australia (Total N = 39)</th>
<th>Brazil (Total N = 27)</th>
<th>China (Total N = 45)</th>
<th>United States (Total N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>11. When you think about public health programs that have ended, what are the most common reasons for programs ending?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Program was never evaluated</td>
<td>9 (23.1)</td>
<td>9 (33.3)</td>
<td>3 (6.7)</td>
<td>9 (16.7)</td>
</tr>
<tr>
<td>Program was evaluated but did not demonstrate impact</td>
<td>12 (30.8)</td>
<td>8 (29.6)</td>
<td>16 (55.6)</td>
<td>7 (13.0)</td>
</tr>
<tr>
<td>Opposition/lack of support from leaders in my agency</td>
<td>7 (17.9)</td>
<td>10 (37.0)</td>
<td>8 (17.8)</td>
<td>6 (11.1)</td>
</tr>
<tr>
<td>Opposition/lack of support from the general public</td>
<td>1 (2.6)</td>
<td>6 (22.2)</td>
<td>20 (44.4)</td>
<td>6 (11.1)</td>
</tr>
<tr>
<td>Opposition/lack of support from policy makers</td>
<td>10 (25.6)</td>
<td>8 (29.6)</td>
<td>10 (22.2)</td>
<td>10 (18.5)</td>
</tr>
<tr>
<td>Funding diverted to a higher priority program</td>
<td>11 (28.2)</td>
<td>13 (48.1)</td>
<td>12 (26.7)</td>
<td>20 (37.0)</td>
</tr>
<tr>
<td>Grant funding ended</td>
<td>25 (64.1)</td>
<td>14 (51.9)</td>
<td>12 (26.7)</td>
<td>46 (85.2)</td>
</tr>
<tr>
<td>Change in political leadership</td>
<td>17 (43.6)</td>
<td>15 (55.6)</td>
<td>3 (6.7)</td>
<td>4 (7.4)</td>
</tr>
<tr>
<td>Insurance funding/coverage ended</td>
<td>1 (2.6)</td>
<td>4 (14.8)</td>
<td>0 (0.0)</td>
<td>6 (11.1)</td>
</tr>
<tr>
<td>Program was adopted or continued by other organizations</td>
<td>0 (0.0)</td>
<td>1 (3.7)</td>
<td>0 (0.0)</td>
<td>8 (14.8)</td>
</tr>
<tr>
<td>A program champion departed</td>
<td>9 (23.1)</td>
<td>9 (33.3)</td>
<td>3 (6.7)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Program was not evidence-based</td>
<td>1 (2.6)</td>
<td>7 (25.9)</td>
<td>5 (11.1)</td>
<td>3 (5.6)</td>
</tr>
<tr>
<td>Program was expensive</td>
<td>1 (2.6)</td>
<td>4 (14.8)</td>
<td>8 (17.8)</td>
<td>11 (20.4)</td>
</tr>
<tr>
<td>Program was challenging to maintain</td>
<td>2 (5.1)</td>
<td>4 (14.8)</td>
<td>24 (53.3)</td>
<td>2 (3.7)</td>
</tr>
<tr>
<td>Other, please specify _______</td>
<td>2 (5.1)</td>
<td>0 (0.0)</td>
<td>1 (2.2)</td>
<td>2 (3.7)</td>
</tr>
<tr>
<td>I do not know</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>4 (8.9)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Not applicable</td>
<td>1 (2.6)</td>
<td>1 (3.7)</td>
<td>0 (0.0)</td>
<td>5 (9.3)</td>
</tr>
<tr>
<td>12. In your opinion, how often do programs continue that should have ended?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1 (2.6)</td>
<td>4 (14.8)</td>
<td>19 (42.2)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Sometimes</td>
<td>20 (51.3)</td>
<td>21 (77.8)</td>
<td>2 (4.4)</td>
<td>35 (64.8)</td>
</tr>
<tr>
<td>Often</td>
<td>10 (25.6)</td>
<td>2 (7.4)</td>
<td>22 (48.9)</td>
<td>13 (24.1)</td>
</tr>
<tr>
<td>13. When you think about public health programs that continued that should have ended, what are the most common reasons for their continuation?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Program was never evaluated</td>
<td>12 (30.8)</td>
<td>7 (25.9)</td>
<td>6 (13.3)</td>
<td>9 (16.7)</td>
</tr>
<tr>
<td>Sustained support from leaders in your agency</td>
<td>13 (33.3)</td>
<td>6 (22.2)</td>
<td>14 (31.1)</td>
<td>16 (29.6)</td>
</tr>
<tr>
<td>Sustained support from the general public</td>
<td>6 (15.4)</td>
<td>5 (18.5)</td>
<td>17 (37.8)</td>
<td>8 (14.8)</td>
</tr>
<tr>
<td>Sustained support from policymakers</td>
<td>11 (28.2)</td>
<td>12 (44.4)</td>
<td>15 (33.3)</td>
<td>21 (38.9)</td>
</tr>
<tr>
<td>Prohibitive costs of starting something new</td>
<td>9 (23.1)</td>
<td>4 (14.8)</td>
<td>5 (11.1)</td>
<td>5 (9.3)</td>
</tr>
<tr>
<td>Absence of alternative options</td>
<td>9 (23.1)</td>
<td>7 (25.9)</td>
<td>13 (28.9)</td>
<td>9 (16.7)</td>
</tr>
<tr>
<td>Sustained funding</td>
<td>7 (17.9)</td>
<td>12 (44.4)</td>
<td>14 (31.1)</td>
<td>26 (48.1)</td>
</tr>
<tr>
<td>Presence of a program champion</td>
<td>12 (30.8)</td>
<td>7 (25.9)</td>
<td>5 (11.1)</td>
<td>13 (24.1)</td>
</tr>
<tr>
<td>Program was considered evidence-based</td>
<td>4 (10.3)</td>
<td>2 (7.4)</td>
<td>9 (20.0)</td>
<td>5 (9.3)</td>
</tr>
<tr>
<td>Program was low-cost</td>
<td>9 (23.1)</td>
<td>6 (22.2)</td>
<td>5 (11.1)</td>
<td>11 (20.4)</td>
</tr>
<tr>
<td>Program was easy to maintain</td>
<td>10 (25.6)</td>
<td>6 (22.2)</td>
<td>12 (26.7)</td>
<td>15 (27.8)</td>
</tr>
<tr>
<td>Other, please specify _______</td>
<td>4 (10.3)</td>
<td>1 (3.7)</td>
<td>0 (0.0)</td>
<td>4 (7.4)</td>
</tr>
<tr>
<td>I do not know</td>
<td>0 (0.0)</td>
<td>3 (11.1)</td>
<td>5 (11.1)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Not applicable</td>
<td>2 (5.1)</td>
<td>0 (0.0)</td>
<td>3 (6.7)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>14. Which of the following are personal barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not being an expert on relevant issues</td>
<td>10 (25.6)</td>
<td>5 (18.5)</td>
<td>29 (64.4)</td>
<td>12 (22.2)</td>
</tr>
<tr>
<td>Lack of confidence in finding data and statistics</td>
<td>7 (17.9)</td>
<td>1 (3.7)</td>
<td>10 (22.2)</td>
<td>6 (11.1)</td>
</tr>
<tr>
<td>Lack of skills to develop evidence-based interventions</td>
<td>6 (15.4)</td>
<td>6 (22.2)</td>
<td>18 (40.0)</td>
<td>8 (14.8)</td>
</tr>
</tbody>
</table>

(Continued)
### TABLE 2 | Continued

<table>
<thead>
<tr>
<th>Question and response options</th>
<th>Australia (Total N = 39)</th>
<th>Brazil (Total N = 27)</th>
<th>China (Total N = 45)</th>
<th>United States (Total N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>Lack of confidence in carrying out evidence-based interventions</td>
<td>3 (7.7)</td>
<td>1 (3.7)</td>
<td>7 (15.6)</td>
<td>3 (5.6)</td>
</tr>
<tr>
<td>Lack of decision-making authority</td>
<td>23 (59.0)</td>
<td>8 (29.6)</td>
<td>20 (44.4)</td>
<td>15 (27.8)</td>
</tr>
<tr>
<td>Low value of evidence-based approaches</td>
<td>5 (12.8)</td>
<td>13 (48.1)</td>
<td>0 (0.0)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Workload is too heavy/not enough time</td>
<td>19 (48.7)</td>
<td>5 (18.5)</td>
<td>19 (42.2)</td>
<td>33 (61.1)</td>
</tr>
<tr>
<td>Overwhelmed by task</td>
<td>5 (12.8)</td>
<td>6 (22.2)</td>
<td>6 (13.3)</td>
<td>11 (20.4)</td>
</tr>
<tr>
<td>Other</td>
<td>8 (20.5)</td>
<td>1 (3.7)</td>
<td>1 (2.2)</td>
<td>12 (22.2)</td>
</tr>
<tr>
<td>None</td>
<td>2 (5.1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>5 (9.3)</td>
</tr>
</tbody>
</table>

15. Which of the following are agency-level barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions?

- Poor understanding of evidence-based approaches: 6 (15.4) Australia, 5 (18.5) Brazil, 4 (8.9) China, 12 (22.2) United States
- Culture/climate is not supportive of change/new ideas: 14 (35.9) Australia, 3 (11.1) Brazil, 1 (2.2) China, 20 (37.0) United States
- No existing policies to support evidence-based approaches: 5 (12.8) Australia, 4 (14.8) Brazil, 14 (31.1) China, 9 (16.7) United States
- Agency does not provide training in evidence-based approaches: 9 (23.1) Australia, 10 (37.0) Brazil, 5 (11.1) China, 10 (18.5) United States
- Staff/leaders lack formal training in evidence-based approaches: 12 (30.8) Australia, 7 (25.9) Brazil, 13 (28.9) China, 14 (25.9) United States
- Lack of access to resources (e.g., computer, Internet): 4 (10.3) Australia, 3 (11.1) Brazil, 14 (31.1) China, 3 (5.6) United States
- Not enough funding: 22 (56.4) Australia, 13 (48.1) Brazil, 13 (28.9) China, 40 (74.1) United States
- Low priority placed on chronic disease prevention: 5 (12.8) Australia, 5 (18.5) Brazil, 4 (8.9) China, 12 (22.2) United States
- No systems to ensure interventions are evidence-based: 8 (20.5) Australia, 16 (59.3) Brazil, 11 (24.4) China, 15 (27.8) United States
- Not enough staff: 11 (28.2) Australia, 4 (14.8) Brazil, 17 (37.8) China, 30 (55.6) United States
- Beliefs that evidence-based interventions are too difficult to implement/sustain: 4 (10.3) Australia, 2 (7.4) Brazil, 4 (8.9) China, 4 (7.4) United States
- Other: 6 (15.4) Australia, 0 (0.0) Brazil, 0 (0.0) China, 2 (3.7) United States
- None: 3 (7.7) Australia, 0 (0.0) Brazil, 3 (6.7) China, 3 (5.6) United States

16. Which of the following are community-level barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions?

- Lack of access to repositories/databases of scientific studies: 7 (17.9) Australia, 4 (14.8) Brazil, 37 (82.2) China, 7 (13.0) United States
- Lack of partnership between agency and community: 13 (33.3) Australia, 4 (14.8) Brazil, 6 (13.3) China, 13 (24.1) United States
- Community members' needs compete with evidence-based recommendations: 22 (56.4) Australia, 8 (29.6) Brazil, 20 (44.4) China, 30 (55.6) United States
- Catering to preferences of funders: 17 (43.6) Australia, 3 (11.1) Brazil, – China, 25 (46.3) United States
- Low priority placed on chronic disease prevention: 11 (28.2) Australia, 14 (51.9) Brazil, 7 (15.6) China, 15 (27.8) United States
- Other: 4 (10.3) Australia, 1 (3.7) Brazil, 0 (0.0) China, 6 (11.1) United States
- None: 1 (2.6) Australia, 7 (25.9) Brazil, 3 (6.7) China, 4 (7.4) United States

17. Which of the following are sociocultural barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions?

- Distrust of scientific data in the populations served: 2 (5.1) Australia, 5 (18.5) Brazil, 7 (15.6) China, 13 (24.1) United States
- Community cultural practices conflict with evidence-based recommendations: 20 (51.3) Australia, 9 (33.3) Brazil, 13 (28.9) China, 19 (35.2) United States
- Not enough relevant evidence for populations served: 18 (46.2) Australia, 5 (18.5) Brazil, 25 (55.6) China, 14 (25.9) United States
- Serving a rural setting where data are lacking: 15 (38.5) Australia, 3 (11.1) Brazil, – China, 34 (63.0) United States
- Serving a highly disadvantaged population: 18 (46.2) Australia, 7 (25.9) Brazil, 10 (22.2) China, 21 (38.9) United States
- Serving a population that speaks a language different from the majority: 7 (17.9) Australia, 1 (3.7) Brazil, – China, 8 (14.8) United States
- Evidence is presented in a language I do not understand: 0 (0.0) Australia, 1 (3.7) Brazil, 6 (13.3) China, 2 (3.7) United States
- Other: 1 (2.6) Australia, 8 (29.6) Brazil, 1 (2.2) China, 2 (3.7) United States
- None: 3 (7.7) Australia, 2 (7.4) Brazil, 4 (8.9) China, 4 (7.4) United States

(Continued)
TABLE 2 | Continued

<table>
<thead>
<tr>
<th>Question and response options</th>
<th>Australia (Total N = 39)</th>
<th>Brazil (Total N = 27)</th>
<th>China (Total N = 45)</th>
<th>United States (Total N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18. Which of the following are political/economic barriers that make it harder for you to select and implement evidence-based chronic disease prevention interventions?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political leaders not providing enough support</td>
<td>24 (61.5)</td>
<td>10 (37.0)</td>
<td>26 (57.8)</td>
<td>21 (38.9)</td>
</tr>
<tr>
<td>Funding changes that occur with changes in political leadership</td>
<td>33 (84.6)</td>
<td>12 (44.4)</td>
<td>13 (28.9)</td>
<td>31 (57.4)</td>
</tr>
<tr>
<td>Political climate conflicts with evidence-based chronic disease prevention recommendations</td>
<td>21 (53.8)</td>
<td>2 (7.4)</td>
<td>16 (35.6)</td>
<td>25 (46.3)</td>
</tr>
<tr>
<td>Health care system does not support evidence-based chronic disease prevention</td>
<td>15 (38.5)</td>
<td>5 (18.5)</td>
<td>3 (6.7)</td>
<td>15 (27.8)</td>
</tr>
<tr>
<td>Other</td>
<td>5 (12.8)</td>
<td>3 (11.1)</td>
<td>0 (0.0)</td>
<td>2 (3.7)</td>
</tr>
<tr>
<td>None</td>
<td>1 (2.6)</td>
<td>0 (0.0)</td>
<td>12 (26.7)</td>
<td>5 (9.3)</td>
</tr>
<tr>
<td>19. For which of the following skills would you like additional technical support or training:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prioritizing program and policy options</td>
<td>22 (56.4)</td>
<td>14 (51.9)</td>
<td>15 (33.3)</td>
<td>21 (38.9)</td>
</tr>
<tr>
<td>Quantifying the public health issue using descriptive epidemiology (e.g., concepts of person, place, time)</td>
<td>18 (46.2)</td>
<td>4 (14.8)</td>
<td>14 (31.1)</td>
<td>20 (37.0)</td>
</tr>
<tr>
<td>Using quantitative evaluation approaches (e.g., surveillance or surveys)</td>
<td>15 (38.5)</td>
<td>6 (22.2)</td>
<td>11 (24.4)</td>
<td>19 (35.2)</td>
</tr>
<tr>
<td>Using qualitative evaluation approaches (e.g., focus groups, key informant interviews)</td>
<td>13 (33.3)</td>
<td>13 (48.1)</td>
<td>19 (42.2)</td>
<td>17 (31.5)</td>
</tr>
<tr>
<td>Developing an action plan for achieving goals</td>
<td>12 (30.8)</td>
<td>17 (63.0)</td>
<td>26 (57.8)</td>
<td>27 (50.0)</td>
</tr>
<tr>
<td>Defining the health issue according to the community’s needs and assets</td>
<td>16 (41.0)</td>
<td>16 (59.3)</td>
<td>12 (26.7)</td>
<td>29 (53.7)</td>
</tr>
<tr>
<td>Adapting interventions for different communities and settings</td>
<td>21 (53.8)</td>
<td>12 (44.4)</td>
<td>16 (35.6)</td>
<td>22 (40.7)</td>
</tr>
<tr>
<td>Using economic data in the decision making process</td>
<td>20 (51.3)</td>
<td>8 (29.6)</td>
<td>15 (33.3)</td>
<td>23 (42.6)</td>
</tr>
<tr>
<td>Communicating research to policy makers</td>
<td>1 (2.6)</td>
<td>0 (0.0)</td>
<td>1 (2.2)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (2.6)</td>
<td>7 (25.9)</td>
<td>3 (6.7)</td>
<td>2 (3.7)</td>
</tr>
</tbody>
</table>

*This item was not applicable and not included in the survey for respondents in China.*

among chronic disease prevention practitioners in Australia, Brazil, China, and the United States, ideally in representative samples (45).

There was low prevalence ($N < 5$) for many response options and the items with low prevalence varied by country. According to Sim and Wright, low prevalence has stifling effects on Cohen’s kappa coefficients, but inflating effects on percentage agreement (46). Low prevalence likely contributed to the low kappa coefficients and comparatively higher percentage agreement found in this study. A larger sample of practitioners across all four countries with more diversity of experiences may improve the variability of responses and the accuracy of reliability findings. Response items with low prevalence of endorsements may also reflect response items that are less applicable to practitioners’ experiences in that particular country. Use of this survey tool in a larger, randomly selected sample of chronic disease practitioners in each country would clarify this conjecture.

**Strengths and Limitations**

This study responds well to a U.S. federal report that called for additional research focused on the experiences and perspectives of key stakeholders in evidence-based intervention delivery, in order to better facilitate the sustainability of interventions (47). The questions within this survey tool reflect critical contextual factors based on the literature, qualitative interviews of public health practitioners, and expert review (2, 5, 6). This survey tool allows researchers to proceed with research on the D&I of EBCDP interventions on a more global scale than was previously available. To our knowledge, this is the first study of its kind that used common methods across four countries. The research team had particular trouble recruiting retest respondents in Brazil due to significant political unrest that affected public health practitioners at the time of the request (39, 40). This contributed to the longer duration between test and retest and the smaller sample from Brazil compared with the other three countries. Additionally, his survey tool demonstrated lower reliability of responses among practitioners from Brazil and China compared with those from Australia and the United States. Lastly, a convenience sampling approach was carried out in some of the countries to recruit chronic disease prevention practitioners serving local or regional jurisdictions. Such a sampling method introduces potential selection bias and is unlikely to produce representative samples of all chronic disease prevention practitioners in each country. However, the intention...
TABLE 3 | Test-retest percent agreement and reliability coefficients by question and country (N = 165).

<table>
<thead>
<tr>
<th>Questions</th>
<th>Australia (N = 39)</th>
<th>Brazil (N = 27)</th>
<th>China (N = 45)</th>
<th>United States (N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (^a)</td>
<td>ICC (^b,c)</td>
<td>%</td>
<td>ICC</td>
</tr>
<tr>
<td>Personal knowledge of evidence-based processes</td>
<td>0.564</td>
<td>0.570</td>
<td>0.667</td>
<td>0.003</td>
</tr>
<tr>
<td>Personal use of repositories to find evidence-based interventions</td>
<td>0.583</td>
<td>0.544</td>
<td>0.630</td>
<td>0.594</td>
</tr>
<tr>
<td>Workplace staff use of repositories to find evidence-based interventions</td>
<td>0.828</td>
<td>0.762</td>
<td>0.539</td>
<td>0.264</td>
</tr>
<tr>
<td>Percentage workplace programs are evidence-based</td>
<td>−</td>
<td>0.731</td>
<td>−</td>
<td>0.566</td>
</tr>
<tr>
<td>Workplace staff use of quality improvement processes</td>
<td>0.581</td>
<td>0.422</td>
<td>0.539</td>
<td>0.423</td>
</tr>
<tr>
<td>Frequency that programs end that should have continued</td>
<td>0.714</td>
<td>0.297</td>
<td>0.444</td>
<td>0.219</td>
</tr>
<tr>
<td>Frequency that programs continue that should have ended</td>
<td>0.828</td>
<td>0.569</td>
<td>0.682</td>
<td>0.116</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Questions</th>
<th>Mean %</th>
<th>Mean Kappa (^d)</th>
<th>Mean %</th>
<th>Mean Kappa (^d)</th>
<th>Mean %</th>
<th>Mean Kappa (^d)</th>
<th>Mean %</th>
<th>Mean Kappa (^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Important factors in decision-making related to program planning and implementation, policy development, or funding</td>
<td>0.811</td>
<td>0.305</td>
<td>0.768</td>
<td>0.416</td>
<td>0.739</td>
<td>0.311</td>
<td>0.772</td>
<td>0.361</td>
</tr>
<tr>
<td>Avenues used to learn about evidence-based chronic disease prevention interventions</td>
<td>0.774</td>
<td>0.402</td>
<td>0.702</td>
<td>0.297</td>
<td>0.804</td>
<td>0.225</td>
<td>0.786</td>
<td>0.408</td>
</tr>
<tr>
<td>Avenues for which additional access is needed</td>
<td>0.810</td>
<td>0.214</td>
<td>0.763</td>
<td>0.154</td>
<td>0.800</td>
<td>0.252</td>
<td>0.778</td>
<td>0.177</td>
</tr>
<tr>
<td>Most common reasons for program termination</td>
<td>0.824</td>
<td>0.203</td>
<td>0.693</td>
<td>0.203</td>
<td>0.745</td>
<td>0.179</td>
<td>0.840</td>
<td>0.289</td>
</tr>
<tr>
<td>Most common reasons for program continuation</td>
<td>0.758</td>
<td>0.258</td>
<td>0.684</td>
<td>0.219</td>
<td>0.733</td>
<td>0.133</td>
<td>0.754</td>
<td>0.249</td>
</tr>
<tr>
<td>Personal barriers to selecting and implementing evidence-based chronic disease prevention interventions</td>
<td>0.823</td>
<td>0.379</td>
<td>0.810</td>
<td>0.308</td>
<td>0.749</td>
<td>0.336</td>
<td>0.857</td>
<td>0.376</td>
</tr>
<tr>
<td>Organizational barriers to selecting and implementing evidence-based chronic disease prevention interventions</td>
<td>0.783</td>
<td>0.279</td>
<td>0.790</td>
<td>0.362</td>
<td>0.794</td>
<td>0.260</td>
<td>0.838</td>
<td>0.462</td>
</tr>
<tr>
<td>Community-level barriers to selecting and implementing evidence-based chronic disease prevention interventions</td>
<td>0.747</td>
<td>0.235</td>
<td>0.763</td>
<td>0.337</td>
<td>0.836</td>
<td>0.367</td>
<td>0.798</td>
<td>0.387</td>
</tr>
<tr>
<td>Sociocultural barriers to selecting and implementing evidence-based chronic disease prevention interventions</td>
<td>0.756</td>
<td>0.286</td>
<td>0.806</td>
<td>0.353</td>
<td>0.808</td>
<td>0.433</td>
<td>0.848</td>
<td>0.476</td>
</tr>
<tr>
<td>Political/economic barriers to selecting and implementing evidence-based chronic disease prevention interventions</td>
<td>0.763</td>
<td>0.204</td>
<td>0.778</td>
<td>0.391</td>
<td>0.747</td>
<td>0.267</td>
<td>0.782</td>
<td>0.329</td>
</tr>
<tr>
<td>Skills for which additional technical support or training is needed</td>
<td>0.764</td>
<td>0.466</td>
<td>0.619</td>
<td>0.191</td>
<td>0.687</td>
<td>0.257</td>
<td>0.736</td>
<td>0.273</td>
</tr>
</tbody>
</table>

\(^a\) Percent agreement. \(^b\) ICC, Intraclass correlation coefficient. \(^c\) Survey questions with ordinal response options were assessed using ICC. \(^d\) Survey questions with a list of response options had each response option dichotomized into selected or not selected, then assessed using Cohen’s kappa, and the mean kappa for each set of response options is reported.

of the present study was not to test hypotheses or provide prevalence estimates, which would have required using methods to address sampling error (46). Acknowledging these limitations of the sampling approach, the researcher team ensured that the selected sample included practitioners from various regions of each country, and provided distributions of all survey responses as well as demographic characteristics of the sample.

CONCLUSION

This survey tool allows cross-country data collection that can contribute toward an improved understanding of the contextual factors that public health practitioners in Australia, Brazil, China, and the United States face in their daily chronic disease prevention work. This understanding is necessary for the creation of multi-level strategies and policies that promote evidence-based decision-making and effective prevention of chronic diseases on a global scale.

ETHICS STATEMENT

This study was carried out in accordance with the committee responsible for human experimentation (institutional and national) and with the World Medical Association’s Declaration of Helsinki with informed consent from all subjects. After reading the electronic informed consent document, subjects indicated their consent by selecting a radial button at the bottom of the informed consent document that read, I consent to participate in this research study. Additional written documentation of consent was waived and the protocol was approved by The University
of Melbourne Human Ethics Committee, Pontificia Universidade Catolica do Parana Research Ethics Committee, The Hong Kong Polytechnic University Human Ethics Committee of the Faculty of Health and Social Science, and Washington University in St. Louis Institutional Review Board. Reasons for waived written documentation of consent: Electronic documentation of informed consent was deemed sufficient for this study because of the non-sensitive nature of the questions and the participants’ locations in four different countries. The following groups agreed on this decision: The University of Melbourne Human Ethics Committee, Pontificia Universidade Catolica do Parana Research Ethics Committee, The Hong Kong Polytechnic University Human Ethics Committee of the Faculty of Health and Social Science, and Washington University in St. Louis Institutional Review Board.

REFERENCES

AUTHOR CONTRIBUTIONS
EB contributed to the conception and design of the study, interpretation of data, and drafting of the full manuscript. XY and AdR contributed to the analysis and interpretation of data and drafting of the Statistical Analyses. KS and RB contributed to the conception and design of the study, interpretation of data, and drafting of the Discussion. ZW, PS, TP, RA, and RR contributed to the conception and design of the study. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING
This work was supported by the National Cancer Institute of the National Institutes of Health (1R21CA179932-01A1).


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Budd, Ying, Stamatakis, deRuyter, Wang, Sung, Pettman, Armstrong, Reis and Brownson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Organizational Readiness Tools for Global Health Intervention: A Review

James W. Dearing*
Michigan State University, East Lansing, MI, United States

The ability of non-governmental organizations, government agencies, and corporations to deliver and support the availability and use of interventions for improved global public health depends on their readiness to do so. Yet readiness has proven to be a rather fluid concept in global public health, perhaps due to its multidimensional nature and because scholars and practitioners have applied the concept at different levels such as the individual, organization, and community. This review concerns 30 publicly available tools created for the purpose of organizational readiness assessment in order to carry out global public health objectives. Results suggest that these tools assess organizational capacity in the absence of measuring organizational motivation, thus overlooking a key aspect of organizational readiness. Moreover, the tools reviewed are mostly untested by their developers to establish whether the tools do, in fact, measure capacity. These results suggest opportunities for implementation science researchers.

Keywords: organizational readiness tools, global public health, organizational capacity, organizational motivation, implementation science, scale up

Despite a common emphasis on the development of effective global health interventions, the greatest contemporary challenges in improving the health of populations rest with the delivery and utilization of interventions. Delivery relies on many human factors—such as communication, coordination, training, leadership and management, logistics, transportation, storage, and community outreach and behavioral campaigns—that function both independently of and interdependently with technical systems. Delivery is necessarily reliant on systems, comprised of different types of organizations, their histories, and current ways of working together. Strengthening these systems and the organizations that comprise them represents a global health priority.

In this article, I review the state of applied tools for assessing organizational readiness for global health intervention, and suggest how they might be improved through research and evaluation.

READINESS MEANS MOTIVATION COMBINED WITH CAPACITY

The term "readiness" has often meant a psychological state (if measured at organizational and community levels, this represents a shared "state") of commitment to a particular course of action. For example, individual readiness can refer to a person’s resolve to stop smoking, organizational readiness can be a shared belief by hospital staff that hospital acquired infections are unacceptably
With a systematic search of published and unpublished (gray) literature for tools (including decision aids and instruments such as questionnaires) that were designed to provide information about the capacity and motivation of organizations involved in global health.

### INCLUSION CRITERIA AND METHODS

For this review, publically available resources must have had:

- Addressed organizational capacity and/or motivation (tools that assessed general organizational assessment or performance were not included)
- Been relevant to the objectives of global health interventions (resources primarily assessing or addressing system- or individual-level factors were not included)
- Been created for organizations working in global health or international development in low-resource countries
- Contained a tool, instrument, or decision aid that:
  - Facilitates decision-making
  - Provides recommended measures or operationalized frameworks/questions to assess capacity and/or motivation
  - Enables quantitative or qualitative measurement of factors

We defined factors as attributes thought to contribute to organizational motivation and capacity to perform a service or function. A tool, instrument, or decision-aid was defined as a published collection of measures or factors meant to assist individuals in assessing the capacity or motivation of an organization.

Resources were identified using a systematic search of computerized databases (OVID, Web of Science, Academic Search Premier, CSA Sociological Abstracts), in-depth web-based searches, and bibliographic snowballing and back referencing. These search strategies were supplemented with targeted searches of relevant organizations, including World Bank Institute (WBI), the United Nations Development Programme (UNDP), John Snow International (JSI), United States Agency for International Development (USAID), and the UK Department for International Development (DFID), among others.

Our search identified 141 potentially relevant tools. An initial review was conducted by at least two members of the research team to assess each tool against the inclusion criteria. This process was done by reviewing the web sites for each tool and then the instructions and items specific to each tool. Each tool was assessed against each inclusion criterion. Tools were independently reviewed by two trained coders and coded for a range of variables. Differences of opinion were resolved through discussion and, in instances of continued disagreement, by the project manager who had trained the coders.

### RESULTS

Thirty tools met the inclusion criteria. These 30 tools are included in this analysis (see Table 1).

The 30 tools are of different types (decision trees, questionnaires, checklists, matrices, etc.) and formats (paper, mobile, web-based, etc.) as listed in Box 1. Tools had a mean number
of 60 questions or items grouped into a mean number of 13 factors.

All tools addressed capacity; none addressed motivation. Each tool was coded for all factors it addressed and/or measured. Our team inductively developed a composite matrix of the capacity factors represented in the tools. Coders had been trained to familiarize themselves with barriers to or facilitators for the scale up of global health interventions in low-income countries (14–18). We then conducted an iterative analysis to identify those domains and factors most commonly addressed in the 30 tools. We grouped the factors into five domains associated with organizational capacity: (1) External Environment; (2) Organizational Attributes; (3) Management and Governance Capacity; (4) Collaboration; and (5) Organizational Performance. Each domain contains between two and eight factors (see Table 2). Analysis was completed based on totals by domain and factor, and by tool.

The tools reviewed here assessed capacities of organizations by allowing users to enter qualitative and/or quantitative data, derived from expert judgment, interviews with staff and stakeholders, document review, workshops, and observation. Fourteen tools allowed both qualitative and quantitative data, 11 allowed only quantitative input, generally in the form of ordinal scales, and 5 allowed only qualitative input. Eleven tools provided scores or ratings by capacity factor and/or a composite capacity score or rating. A few tools also provide graphic output to provide

### Table 1 | Organization capacity assessment tools (N = 30).

<table>
<thead>
<tr>
<th>Tool name</th>
<th>Developed by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessing Management Capacity Among Non-governmental Organizations</td>
<td>CARE International</td>
</tr>
<tr>
<td>Capacity Assessment Framework</td>
<td>United Nations Development Programme (UNDP)</td>
</tr>
<tr>
<td>Capacity Building in Training-Dimensions and Indicators</td>
<td>PRIME/INTRAH</td>
</tr>
<tr>
<td>Capacity Development Results Framework (CDRF)</td>
<td>World Bank Institute (WBI)</td>
</tr>
<tr>
<td>Discussion-Oriented Organizational Self-Assessment (DOSA)</td>
<td>USAID Center for Development Information and Evaluation</td>
</tr>
<tr>
<td>Dynamic Participatory Institutional Diagnosis (DPID)</td>
<td>Senegal PVO/NGO, New TransCentury Foundation, Yrawah International</td>
</tr>
<tr>
<td>How to Assess NGO Capacity?</td>
<td>Norwegian Missionary Council Office for Development Cooperation</td>
</tr>
<tr>
<td>Institutional Development Framework (IDF)</td>
<td>Management Systems International (MSI)</td>
</tr>
<tr>
<td>Institutional Self Reliance (ISR)</td>
<td>Research Triangle Institute for UNDP</td>
</tr>
<tr>
<td>Management and Organizational Stability Tool (MOSIT)</td>
<td>Management Sciences for Health</td>
</tr>
<tr>
<td>NGO Capacity Analysis</td>
<td>International HIV/AIDS Alliance</td>
</tr>
<tr>
<td>NGO Sustainability Index</td>
<td>USAID Office of Democracy and Governance; USAID Bureau for Europe and Eurasia</td>
</tr>
<tr>
<td>Nonprofit Organizational Assessment Tool-Strategic Planning Assessment Tool</td>
<td>Andrew Lewis, University of Wisconsin Extension</td>
</tr>
<tr>
<td>Organizational Capacity Assessment (OCA) Tool</td>
<td>PACT</td>
</tr>
<tr>
<td>Organizational Capacity Assessment Tool</td>
<td>Marguerite Casey Foundation</td>
</tr>
<tr>
<td>Organizational Capacity Audit Tool</td>
<td>GeSCI (Global e-Schools and Communities Initiative)</td>
</tr>
<tr>
<td>Organizational Capacity Indicator (OCI)</td>
<td>Christian Reformed World Relief Committee (CRWRC)</td>
</tr>
<tr>
<td>Organizational Capacity Self-Assessment Tool</td>
<td>Academy of Educational Development (AED), Croatia’s Non-governmental Sector (CroNGO)</td>
</tr>
<tr>
<td>Participatory Organizational Evaluation Tool (POET)</td>
<td>PACT, United Nations Development Programme (UNDP)</td>
</tr>
<tr>
<td>Partner Assessment Form</td>
<td>The Partnering Initiative/International Business Leaders Forum (IBLF)</td>
</tr>
<tr>
<td>Partner Organizational Capacity Assessment: A Tool for Assessing and Building Capacity for Twinning Partnerships for High-Quality Response to HIV/AIDS</td>
<td>Twinning Center</td>
</tr>
<tr>
<td>Partnership Self-Assessment Tool</td>
<td>Center for the Advancement of Collaborative Strategies in Health</td>
</tr>
<tr>
<td>Rapid Organizational Assessment</td>
<td>Universalia</td>
</tr>
<tr>
<td>Simple Capacity Assessment Tool (SCAT)</td>
<td>Education Development Center, PACT</td>
</tr>
</tbody>
</table>

### Table 1 | Continued

<table>
<thead>
<tr>
<th>Tool name</th>
<th>Developed by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool to Assess Site Readiness for Initiating Antiretroviral Therapy (ART) or Capacities for Existing ART Sites</td>
<td>John Snow, Inc. (JSI)</td>
</tr>
<tr>
<td>Training and Technical Assistance Plan (TTAP)</td>
<td>Counterpart International</td>
</tr>
<tr>
<td>UNDP CAPBUILD User’s Guide</td>
<td>United Nations Development Programme (UNDP)</td>
</tr>
<tr>
<td>Unity Foundation Capacity Quotient: A Diagnostic Tool for Benchmarking Capacity</td>
<td>Unity Foundation</td>
</tr>
<tr>
<td>USAID/Madagascar Institutional Capacity Questionnaire</td>
<td>USAID</td>
</tr>
</tbody>
</table>

### Box 1 | Types and formats of tools.

<table>
<thead>
<tr>
<th>Tool types (types can be &gt; 1)</th>
<th>(N) Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-assessment</td>
<td>24</td>
</tr>
<tr>
<td>External-assessment</td>
<td>10</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>15</td>
</tr>
<tr>
<td>Matrix/grid</td>
<td>6</td>
</tr>
<tr>
<td>Guideline</td>
<td>1</td>
</tr>
<tr>
<td>Decision tree</td>
<td>1</td>
</tr>
<tr>
<td>Tool formats</td>
<td></td>
</tr>
<tr>
<td>Paper-based</td>
<td>26</td>
</tr>
<tr>
<td>Web-based</td>
<td>2</td>
</tr>
<tr>
<td>Electronic</td>
<td>2</td>
</tr>
<tr>
<td>Spreadsheets</td>
<td></td>
</tr>
</tbody>
</table>

Continued
TABLE 2 | Categorization of tool domains and factors.

<table>
<thead>
<tr>
<th>Organizational attributes</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Infrastructure</td>
</tr>
<tr>
<td></td>
<td>Internal communication, knowledge management, and organizational learning</td>
</tr>
<tr>
<td></td>
<td>Leadership</td>
</tr>
<tr>
<td></td>
<td>Mission and vision (mission, strategy, organizational fit)</td>
</tr>
<tr>
<td>Management and governance</td>
<td>Adaptive capacity</td>
</tr>
<tr>
<td></td>
<td>Administration and organizational structure</td>
</tr>
<tr>
<td></td>
<td>Financial resource management</td>
</tr>
<tr>
<td></td>
<td>Human resource management</td>
</tr>
<tr>
<td></td>
<td>Measurement and evaluation</td>
</tr>
<tr>
<td>Collaboration</td>
<td>External partnerships and communication</td>
</tr>
<tr>
<td></td>
<td>Stakeholder partnerships</td>
</tr>
<tr>
<td>Organizational performance</td>
<td>Delivery, procurement, and supply</td>
</tr>
<tr>
<td></td>
<td>Institutional sustainability</td>
</tr>
<tr>
<td></td>
<td>Outputs, service, and results</td>
</tr>
<tr>
<td>External environment</td>
<td>Political/legal environment (including advocacy)</td>
</tr>
<tr>
<td></td>
<td>Sociocultural and geographic environment</td>
</tr>
</tbody>
</table>

Box 2 | Top 10 capacity factors.

<table>
<thead>
<tr>
<th>Factor</th>
<th>% of tools that measure this factor (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission and vision</td>
<td>87% (26)</td>
</tr>
<tr>
<td>Financial management</td>
<td>83% (25)</td>
</tr>
<tr>
<td>Financial resources</td>
<td>80% (24)</td>
</tr>
<tr>
<td>Strategic management</td>
<td>80% (24)</td>
</tr>
<tr>
<td>Administration and organizational structure</td>
<td>80% (24)</td>
</tr>
<tr>
<td>Human resources</td>
<td>73% (22)</td>
</tr>
<tr>
<td>External partnership and communication</td>
<td>73% (22)</td>
</tr>
<tr>
<td>Stakeholder partnership</td>
<td>73% (22)</td>
</tr>
<tr>
<td>Human resource management</td>
<td>67% (20)</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>63% (19)</td>
</tr>
</tbody>
</table>

Assessing the readiness of organizations for global health intervention purposes should, according to the literature, involve measurement of both the capacity and the motivation of those organizations to engage in initiatives. Yet this review found only capacity assessment instruments. Measurement of an organization’s motivation or willingness to prioritize and engage in a global health intervention should be included and made available in such tools either in combination with existing capacity assessment tools or as stand-alone instruments.

Organizational capacity assessment tools with relevance for global health intervention measure many of the same domains and factors. This similarity may reflect a tendency by tool developers to employ common frameworks or orientations about what constitutes organizational capacity for global health intervention. It is possible that tool developers used a common evidence base to derive the measures included in the tools. These observed similarities may highlight an emerging consensus and convergence in scope regarding those factors that predict organizational effectiveness in the delivery of global health interventions.

The tools included in this review are focused at the organizational unit of analysis. As more organizational alliances active in efforts to scale-up health impact emerge, the most relevant level of analysis for estimating likelihood of success will be at the inter-organizational system or partnership level, reflecting the necessity that entire supply chains of collaborating and contracted organizations are involved when interventions are to be delivered to millions of people across large geographic areas (19). At this juncture, we did not find any tools that conceptualized and operationalized capacity or motivation measures at the level of inter-organizational systems or partnerships. While it can be expected that, because systems and inter-organizational
partnerships are comprised of organizational actors, existing measures of organization-level capacities should be relevant, it is also the case that systems and partnerships require greater attention to working with heterophilous others (e.g., ministries of health, private organizations, non-governmental organizations, community health outreach workers). This requires coordination and contracting, and working across organizational boundaries where there may not be systems in place to seamlessly support such large-scale initiatives. Capacity assessment tools do not currently reflect the special challenges of this aggregate level of agency. Yet they could by focusing each partner organization’s experiences with the other organizations, for example, and measures focused on identification of complementary skills and resources across organizations.

**DO ORGANIZATIONAL READINESS ASSESSMENT TOOLS WORK?**

Given the current information available, we cannot conclude whether these assessment tools are effective. Many highly reputable organizations have sponsored the development of these tools; perhaps these tools have been broadly and enthusiastically applied in practice to good effect. Yet we were unable to find data for any of the tools regarding validity assessment or utilization evaluation. None specifically presented evidence supporting the inclusion or validity of specific measures. In general, it appears that developers have relied on expert opinion and structural measures—with an unproven relationship to outcomes—when developing indices and measures. We do not know if use of any of these tools is associated with improvement in organizational motivation or capacities, initiative performance, or efficient use of resources. Neither did we find information about how much any of these tools has been used.

The partial convergence of these tools on similar factors related to capacity suggests that these instruments tap into correct constructs. That is, the factors most commonly assessed are likely meaningfully related to organizational capacity. A number of the 21 factors in this review are well reflected in the literature about scaling up in low-income countries (17, 20). However, improvement and refinement of the composition of items contained in these tools is probably possible.

**UTILIZATION CONSIDERATIONS**

Best practices for instrument and heuristic development suggests that tools or decision aids should be actively tested, refined, and improved with plausible potential users during a pre-testing stage conducted prior to release of the tool. Among other things, stakeholder feedback to prototype tools can provide exceedingly valuable insights into format preferences, optimum length, question order effects, response variance, use of graphics, and which types of potential users are best suited to provide different types of information. Though this may have occurred during development of the reviewed tools, the information available to us rarely included detail about such user-facing formative evaluation. Furthermore, stakeholders can differ considerably in their preferences for applied tools. Yet none of the 30 tools reviewed here was available in more than one interface or format.

Some tools reviewed here were developed in the 1990s, prior to remarkable developments in web-based applications. Nevertheless, the paucity of interactive capacity assessment tools for global health stakeholders is striking. It is likely that many users of these tools would find value in an instrument that is able to provide computations, comparative assessments, confidence intervals, or qualitative but tailored feedback based on data entered by the user. Of course, the utility of interactive formats would be limited to users with access to computers or electronic devices with internet connections.

An important implication for utilization lies in the understanding that much of the value in tools such as these lies beyond the data provided by the user and/or as computational or informational output. The process of engaging in use can be quite valuable for purposes of critical reflection, discussion, the development of a shared understanding among stakeholders, and the stimulation of a collective organizational will to improve processes and outputs. This type of enlightenment use has been shown in some research to be of more consequence in terms of learning and organizational improvement than is the instrumental data (the “answers”) themselves (21).

**NEXT STEPS IN TOOL DEVELOPMENT**

A previously published framework used to evaluate measurement systems in public health (22) employed four criteria to assess instruments:

- Clarity of measurement parameters and normative standards
- Balance between structural and process measures
- Evidence for effectiveness
- Specification of an accountable entity

These criteria, when applied globally to currently existing organizational capacity assessment tools, highlight challenges and opportunities for improvement. Many of the tools we reviewed employ clear standards; the best quantitative scales are tied to well-defined categories that reflect the continuum of organizational development. The newer and more robust tools strike a balance between structural and process measures, as well as specification of accountable entities.

This review suggests a need for robust, validated, and user-friendly tools to measure organizational capacity and, we suggest, organizational motivation; taken together, such tools can more fully represent the construct of organizational readiness for global health intervention. Identified strategies for instrument improvement include standardization, evaluation, validation, and application of an evidence base to inform tool development. This evidence base could, in part, be constructed using retrospective case studies of how global health intervention delivery fared to assess whether successes and failures were associated with certain factors. Information could also be gathered about the valence and weighting of those factors. An evidence base could then be applied prospectively, using predictive tests to determine whether tool use affects roll-out or scale-up of global health interventions, and how. These research validation steps, if applied
in tandem with utilization study during formative development, could result in tools that not only work but also work well for users. As the evidence base supporting the identification of core domains, factors, and appropriate methodologies evolves, tools such as these will become more valid, reliable, and useful for the increasingly diverse range of stakeholders involved in global health interventions.

**AUTHOR CONTRIBUTIONS**

The author originated the project and drafted and rewrote the manuscript.

**REFERENCES**


**ACKNOWLEDGMENTS**

The author thanks Lauren K. Krause, Sarah D. Madrid, Quynh A. Le, Heather A. Nuances, and Erica F. Morse for their work on the present study.

**FUNDING**

This work was supported by the Bill & Melinda Gates Foundation. The information provided in this article solely reflects the view of the author and not the views of the Bill & Melinda Gates Foundation.


**Conflict of Interest Statement:** The author has no commercial conflicts of interest regarding any of the assessment tools reviewed.

Copyright © 2018 Dearing. *This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*
Readiness to Change Over Time: Change Commitment and Change Efficacy in a Workplace Health-Promotion Trial


1 Seattle-Denver Center of Innovation for Veteran-Centered and Value-Driven Care, US Department of Veterans Affairs, Seattle, WA, United States, 2 Department of Health Services, School of Public Health, University of Washington, Seattle, WA, United States, 3 Health Promotion Research Center, A CDC Prevention Research Center, Department of Health Services, University of Washington, Seattle, WA, United States

Introduction: Organizational readiness to change may be a key determinant of implementation success and a mediator of the effectiveness of implementation interventions. If organizational readiness can be reliably and validly assessed at the outset of a change initiative, it could be used to assess the effectiveness of implementation-support activities by measuring changes in readiness factors over time.

Methods: We analyzed two waves of readiness-to-change survey data collected as part of a three-arm, randomized controlled trial to implement evidence-based health promotion practices in small worksites in low-wage industries. We measured five readiness factors: context (favorable broader conditions); change valence (valuing health promotion); information assessment (demands and resources to implement health promotion); change commitment (an intention to implement health promotion); and change efficacy (a belief in shared ability to implement health promotion). We expected commitment and efficacy to increase at intervention sites along with their self-reported effort to implement health promotion practices, termed wellness-program effort. We compared means between baseline and 15 months, and between intervention and control sites. We used linear regression to test whether intervention and control sites differed in their change-readiness scores over time.

Results: Only context and change commitment met reliability thresholds. Change commitment declined significantly for both control (−0.39) and interventions sites (−0.29) from baseline to 15 months, while context did not change for either. Only wellness program effort at 15 months, but not at baseline, differed significantly between control and intervention sites (1.20 controls, 2.02 intervention). Regression analyses resulted in two significant differences between intervention and control sites in changes from baseline to 15 months: (1) intervention sites exhibited significantly smaller change in context scores relative to control sites over time and (2) intervention sites exhibited significantly higher changes in wellness program effort relative to control sites.
INTRODUCTION

Organizational readiness to change is the psychological and behavioral preparedness of organizational members tasked with implementation of a new practice, policy, or technology (1). Organizational readiness is thought to be a key determinant of implementation success and a mediator of the effectiveness of implementation interventions (1–3). Readiness is a core construct in several dissemination and implementation frameworks (4–6).

If organizational readiness can be reliably and validly assessed at the outset of a change initiative, measures of readiness could be used prognostically to gain an accurate prediction of the likelihood of change success and diagnostically to identify specific weaknesses or deficits in readiness. If accurately measured, organizational readiness could be used in workplace health promotion efforts to target work sites for dissemination; to diagnose and address worksite-specific deficits in readiness; and to assess the effectiveness of implementation-support activities by measuring changes in readiness factors over time. Accurate organizational readiness could also be considered or intervened upon with implementation-support activities, such as information, training, and marketing materials.

We tested a previously developed survey designed specifically for assessing organizational readiness to implement evidence-based workplace health promotion practices (7). Our objective was to determine if the instrument was sensitive to changes in readiness factors over time and differences in readiness among workplaces participating in a randomized, controlled implementation trial receiving different implementation-support interventions. Our hope was that the readiness measure could ultimately be used in broader dissemination and implementation efforts to identify workplace-specific implementation barriers that can be addressed with implementation-support activities and potentially repeated to determine if support activities have been successful. However, this is only possible if the readiness measure is sensitive to changes in readiness factors over time and sensitive to improvements in readiness resulting from implementation-support activities. The purpose of this paper is to test the readiness measure’s sensitivity to change over time in work sites that attempted to implement new health promotion practices.

MATERIALS AND METHODS

Design

We analyzed two waves of survey data collected as part of a three-arm, randomized, controlled trial testing the effectiveness of HealthLinks, a workplace-health-promotion program (8). HealthLinks was developed in collaboration with the American Cancer Society and the University of Washington. It is tailored to the needs and capacities of small work sites to help them implement evidence-based practices for workplace health promotion. Work sites participating in HealthLinks receive an assessment of their current implementation of evidence-based practices; a tailored recommendations report; tool kits to support implementation of each of the practices; and onsite, telephone, and email assistance from a trained interventionist. As part of the trial, we developed and validated a readiness-to-change survey, with the goal of creating a survey that could be used in subsequent dissemination efforts (7).

HealthLinks aimed to increase the adoption and implementation of 11 evidence-based health promotion practices through provision of materials and onsite implementation assistance. The evidence-based health promotion practices, recommended by CDC’s Community Guide to Preventive Services (9) as compatible with worksites, focused on healthy eating, physical activity, tobacco cessation, and screening for breast, cervical, and colon cancers (Table 1).

HealthLinks enrolled small work sites in six low-wage industries in King County in Washington State. One intervention arm received only the HealthLinks program (Standard HealthLinks), one arm received HealthLinks plus support to form wellness committees (HealthLinks + Wellness Committee), and the third arm was a delayed control group. As part of the study, work sites completed surveys at baseline and 15 months to assess readiness factors and specific implementation efforts to implement health promotion practices (described below). The study protocol and baseline outcomes have been previously published (8).

Conceptual Model

The readiness measures and analysis were guided by Weiner’s theory of organizational readiness to change (Figure 1) (1). It hypothesizes that organizational readiness to change comprises two collective, affective states: change commitment and change efficacy. Change commitment refers to an intention to implement a change that is shared across members of an organization. Change efficacy is defined as organizational members’ shared beliefs in their joint ability to engage in those courses of action necessary to implement a change. Change-related effort, which we hereafter refer to as wellness-program effort, is the collective effort of organizational members to execute a change, and is a function of both change commitment and change efficacy. While beyond the scope of the current analysis, wellness-program effort is expected to predict the actual extent of implementation of workplace wellness programs.

Discussion: Contrary to our hypothesis, change commitment declined significantly at both HealthLinks and control sites, even as wellness-program effort increased significantly at HealthLinks sites. Regression to the mean may explain the decline in change commitment. Future research needs to assess whether baseline commitment is an independent predictor of wellness-program effort or an effect modifier of the HealthLinks intervention.

Keywords: readiness to change, implementation, change commitment, change efficacy, psychometric validation, workplace health promotion
Change commitment and change efficacy are functions of change valence and informational assessment. Change valence is the extent to which members of an organization value the change. Reasons for why the change is valued can vary, and this construct does not assume that all members value it for the same reason, only that there exists a collective belief that the change is significant to the goals of the organization. Informational assessment refers to organizational members' perceptions that the resources available to implement the change (human, financial, material, and informational) are sufficient to the demand.

Change valence and informational assessment are influenced in turn by context, which refers to the broader conditions that affect readiness to change, such as organizational culture, climate, resources, structure, and past experiences with implementing change. Relative to the other constructs in the model, context is not innovation-specific and should be more stable over time.

### Table 1

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Interventions Promoted in HealthLinks and HealthLinks+</th>
</tr>
</thead>
</table>
| Breast, cervical, and colon cancer screening | • Distribute brochures and post posters to educate workers about cancer screening guidelines  
• Provide brief education sessions at the worksite, including benefits of screening and information about costs/insurance coverage  
• Promote the Washington Breast, Cervical, and Colon Health Program to uninsured workers; include information about local providers, screening free of charge, and treatment coverage for those diagnosed with cancer  
• Promote benefits coverage at those worksites with insurance benefits |
| Healthy eating                    | • For worksites that sell food, create policies to offer healthy options, label them, and price them competitively  
• For all worksites, create policies to support offering healthy foods at meetings and events |
| Physical activity                 | • Negotiate discounts at local gyms for workers  
• Post “Use the Stairs” signs  
• Offer ACS Active for Life program, an evidence-based program that offers individual choice of activity and builds social support |
| Tobacco cessation                 | • Promote the Washington State Tobacco Quit Line via brochures and other small media; include information about quit line services  
• Promote benefits coverage at those worksites with insurance coverage for tobacco cessation |

### Setting and Sample

HealthLinks was tested among small workplaces, defined as 20–200 employees, in low-wage industries in King County in Washington State. We selected industries by North American Industry Classification System (NAICS) codes: accommodation and food services; arts, entertainment, and recreation; education; health care and social assistance; retail trade; and other services excluding public administration. We required eligible worksites to have a minimum of 20% of their employees report to a physical site at least once per week; to have been in business for at least 3 years; and worksites could not already have a wellness committees in place. A total of 78 sites were enrolled, with 28 assigned to the HealthLinks arm; 26 assigned to the HealthLinks plus wellness committee arm; and 24 assigned to the delayed controls.

### Data Collection and Measures

This analysis used three measures: company characteristics, organizational readiness scales, and implementation-related efforts, referred to as wellness program effort. Company characteristics included type of industry, number of employees (size), for-profit vs. not-for-profit, proportion of full-time employees, and whether health insurance was offered to employees.

The readiness to change and wellness program-effort scales were previously developed for this study through a multi-stage validation process. First, we identified existing readiness scales that measured constructs in the Weiner readiness model, starting with the organizational readiness to change measure developed by Shea and colleagues (10) based on the Weiner model, as well other readiness to change surveys (11, 12), and a prior wellness-program survey (13). We then conducted think aloud interviews with employers similar to our study sample to evaluate and revise items for comprehension and appropriateness. Finally, we piloted the survey with a sample of 201 small Washington employers in the same industries as HealthLinks (separate from our HealthLinks sample) in order to assess scale reliability and criterion validity. The latter included a path analysis of scales to confirm that associations among the scales conformed to Weiner's theory of organizational readiness. Survey development and validation procedures and findings were reported in detail in a prior paper (7).

Readiness items (Table S1 in Supplementary Material) were scored on 5-point Likert scales (1 = strongly disagree, 5 = strongly agree). The context scale comprised 10 items assessing leadership, management, and opinion leaders’ willingness to trying new things; whether they reward creativity and innovation; whether they promote teambuilding to solve worksite problems; and whether they seek to improve workplace climate.
The information assessment scale comprised five items assessing availability of staff time, financial resources, and employee and leadership champions for wellness programs. The change valence scale comprised four items assessing whether wellness programs would benefit the organization in terms of improving employee health, improving employee recruitment and retention, and reducing employee health-care costs. The change commitment scale comprised five items assessing senior leader, opinion leader, and collective commitment and motivation to start or improve a wellness program. The change efficacy scale comprised four items assessing collective skills, expertise, ability to manage workplace politics, and ability to obtain employee participation, while implementing a wellness program.

Wellness program effort was measured via five questions about implementation activities for wellness programs, such as having written wellness goals, a wellness committee and coordinator, and/or a health promotion or wellness budget. The fifth item, how much time the respondent thought s/he could spend on a wellness program, was not included in the original development of the wellness program effort scale. We added it here because time spent on wellness activities is an additional and concrete indicator of wellness program effort. The time spent on wellness program effort item was a five-point (1–5) Likert-type scale. Yes–no items, initially coded in the data as yes = 1, no = 0, were re-coded yes = 5, no = 1 to align with the scoring of scale items throughout the readiness survey instrument.

Data were collected through surveys conducted in person at baseline and via telephone at 15-month follow-up. With the exception of a section on satisfaction with the HealthLinks program at follow-up, worksites answered identical sets of questions at baseline and follow-up. Surveys were completed by the primary worksite contact for the study, usually, the Human Resource manager, who would be involved in any workplace health promotion efforts. Delayed control sites received the HealthLinks intervention after data collection ended.

Analysis
We examined the means of measures at baseline and 15 months, and among intervention groups, and tested mean differences using a paired t-test. We used a significance level of \( p \leq 0.05 \). We then examined the association between readiness and wellness-program effort score change and intervention groups using linear regression models, adjusting for worksite size (20–49 vs. 50–200), and industry (arts, entertainment, and recreation/education/health care and social assistance v. accommodation and food services/other services excluding public administration/retail trade), which were the blocking variables for trial randomization and have previously been found to be related to workplace health promotion practices (13). Our hypothesis was that change commitment, change efficacy, and wellness-program effort would increase significantly from baseline to 15 months at intervention sites while not changing significantly at control sites. We used the difference scores of the baseline and 15-month surveys as our outcome measures, and there was a single survey respondent per site.

Our initial analyses examined each of the three study arm sites compared to the other two study arm sites, and the two intervention arm sites (HealthLinks and HealthLinks + Wellness Committee) compared to control sites. As we saw few differences between the intervention sites, we focus the results below on the analyses comparing the combined intervention sites to control sites.

Analyses were conducted with STATA version 15 (College Station, TX, USA).

**Results**
All 78 worksites completed baseline surveys; 72 (92.3%) completed follow-up surveys. Our analyses included the 72 worksites with complete baseline and follow-up data. Intervention and control sites did not differ in industry characteristics (Table 2).

Three readiness scales failed to meet reliability thresholds, as measured by Cronbach’s alpha (Table 3): Change valence (Cronbach’s alpha = 0.66 at baseline, 0.67 at 15 months), information assessment (0.64 at baseline, 0.54 at 15 months), and change efficacy (0.52 at baseline, 0.63 at 15 months). Context (0.72 at baseline, 0.79 at 15 months) and change commitment (0.72 at baseline, 0.71 at follow-up) met reliability thresholds. We only report subsequent findings for scales that exhibited reliability. We did not calculate alpha statistics for wellness program effort because four of the five items were dichotomous, which are not suitable for Crohnbach’s alpha.

When assessing the differences between baseline and 15-month scores (Table 4), change commitment declined significantly for both control (−0.39) and interventions sites (−0.29), while context did not change for either control or intervention sites. When examining the change from baseline to 15 months for each intervention arm separately, the sites in the HealthLinks + wellness committee arm did not see a significant difference in change commitment. Wellness program effort, the proximal outcome, increased significantly for intervention sites (0.73) but did not change for control sites.

When assessing the differences between intervention and control sites for each scale and the outcome at each time period (baseline and 15 months), the only significant difference was for wellness program effort at 15 months (1.20 controls, 2.02 intervention, \( p < 0.05 \)) (Table 5).

Regression analyses resulted in two significant differences between intervention and control sites in changes over the 15 months from baseline to follow-up (Table 6). First, the change in context scores from baseline to follow-up was significantly lower for intervention sites relative to control sites. Second, intervention sites exhibited significantly higher changes in wellness program effort relative to control sites. There were no differences between intervention and control sites in the change in change commitment from baseline to follow-up.

In secondary analyses, we evaluated the reliability of scales, following the procedures we used in our original validation study, to determine if scale reliability could be improved by eliminating
items that had an item-rest correlation of 0.20 or lower. This procedure improved scale reliability but did not result in any change in the scales meeting our threshold of 0.70 (results available upon request).

**DISCUSSION**

Contrary to our hypothesis, change commitment declined significantly at Healthlinks sites, even as wellness-program effort increased significantly. One explanation for this apparent incongruity (declining commitment in the face of increasing effort) could be change fatigue: a gradual exhaustion of participants’ motivation over time as a consequence of their sustained change efforts. However, change commitment declined equally at control sites who were not engaged in any change efforts. The more likely explanation is regression to the mean. Sites were recruited over a period of 10 months and, most likely, motivation to engage in workplace health promotion varies randomly over time. Motivation to engage in workplace health promotion almost certainly correlated with interest in participating in the study, and sites that enrolled in the study were probably often randomly waxing in motivation at the time they decided to enroll. The decline in their 15-month scores may just represent a return to something closer to their average motivation or commitment to workplace wellness. We see indirect evidence of regression to the mean from comparing the change commitment scores observed in this study to the scores observed in the cross-sectional survey used in our prior scale-validation study (7): the mean scores on change commitment in that survey was 3.31, nearly identical to the change commitment scores at 15 months.

When analyzing changeover time, we also found that the difference in context scores from baseline to 15 months was significantly smaller at intervention sites relative to controls. The context scale measures attitudes and actions of senior leaders, managers, and opinion leaders related to workplace climate, creativity, innovation, and team-building to solve worksite
TABLE 4 | Change in organizational readiness to change factors between baseline and 15 months.

<table>
<thead>
<tr>
<th></th>
<th>Baseline Mean</th>
<th>Baseline SD</th>
<th>15 months Mean</th>
<th>15 months SD</th>
<th>Difference*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>Control 3.48</td>
<td>0.40</td>
<td>Intervention 3.62</td>
<td>0.46</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.61</td>
<td>0.49</td>
<td>–0.18</td>
</tr>
<tr>
<td>Change valence</td>
<td>Control 3.96</td>
<td>0.42</td>
<td>Intervention 3.95</td>
<td>0.55</td>
<td>–0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.01</td>
<td>0.49</td>
<td>–0.14</td>
</tr>
<tr>
<td>Information assessment</td>
<td>Control 3.73</td>
<td>0.57</td>
<td>Intervention 3.59</td>
<td>0.53</td>
<td>–0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.59</td>
<td>0.61</td>
<td>–0.06</td>
</tr>
<tr>
<td>Change commitment</td>
<td>Control 3.69</td>
<td>0.57</td>
<td>Intervention 3.30</td>
<td>0.59</td>
<td>–0.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.65</td>
<td>0.56</td>
<td>–0.29</td>
</tr>
<tr>
<td>Change efficacy</td>
<td>Control 3.51</td>
<td>0.61</td>
<td>Intervention 3.40</td>
<td>0.60</td>
<td>–0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.46</td>
<td>0.57</td>
<td>–0.07</td>
</tr>
<tr>
<td>Wellness program effort</td>
<td>Control 1.15</td>
<td>0.32</td>
<td>Intervention 1.20</td>
<td>0.43</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.29</td>
<td>0.47</td>
<td>0.73</td>
</tr>
</tbody>
</table>

*Differences in bold italic are mean values from baseline to 15 months that are significant at p-value ≤ 0.05.

Behavioral interventions were designed to test the sensitivity of an organizational readiness-to-change measure on change readiness. First, ours is the only study we are aware of that evaluated the intervention using a large sample of employers similar to the present study sample, found good reliability and criterion validity (7).

Meanwhile, the scales measuring change efficacy, change valence, and informational assessment exhibited poor reliability and, consequently, we cannot draw any conclusions about the sensitivity of these measures to differences over time or among study arms. The poor reliability of these scales is perplexing; our prior validation of the survey, which included concurrent validation using a large sample of employers similar to the present study sample, found good reliability and criterion validity (7).

One explanation for the poor reliability may be a combination of sample size and systematic measurement error. Shevlin and colleagues have used Monte Carlo simulations to show that alpha coefficients are highly sensitive to the combination of sample size and the presence of measurement error, and the differences we found between our validation and trial data are generally within the differences they observed (14). We know our trial sample size was significantly lower (n = 72) than our validation study (n = 201). In addition, we might expect that the validation study (but not the trial) was susceptible to systematic error due to “halo effect,” because the validation study assessed readiness factors concurrently with extent of workplace health promotion practice. Halo effect is a type of inferential bias in which individuals form a general impression of someone or something and infer other qualities from that general impression, e.g., inferring an individual’s leadership qualities from how well one likes the individual (15). Cross-sectional criterion validation, in which we assess the criterion outcome at the same time as we assess readiness factors, is particularly susceptible to halo effect because the respondent already knows the outcome as they respond to readiness factors. Cross-sectional criterion validation is particularly susceptible to halo effect because the respondent already knows the outcome as they respond to readiness factors concurrently with extent of workplace health promotion practice. Halo effect is a type of inferential bias in which individuals form a general impression of someone or something and infer other qualities from that general impression, e.g., inferring an individual’s leadership qualities from how well one likes the individual (15).

This article makes several contributions to the broader literature on change readiness. First, ours is the only study we are aware of to test the sensitivity of an organizational readiness-to-change measure to changes over time, and the findings ran contrary to our hypotheses. Our study used experimental manipulation that successfully induced greater implementation efforts among intervention sites, creating a scenario in which we had a strong theoretical rationale for expecting significantly greater change.
commitment and change efficacy over time at intervention sites relative to control sites. Yet, we observed no differences between intervention and control sites in commitment, and contrary to expectation, observed declining commitment over time among all sites. This is important because change commitment and change efficacy and related affective constructs such as intention and motivation, are central to most organizational readiness to change measures, and the vast majority of empirical work in this area has historically been cross-sectional or using other designs that are susceptible to bias, e.g., case studies, one-group pretest, posttest (16). We would like to see this experiment replicated in other health promotion contexts, and other implementation fields, to see if similar or different associations are found. That would help advance our underlying conceptual understanding of collective readiness as a prerequisite for effective organizational change.

Second, commitment is core to many implementation models as a mediator of implementation activities and implementation outcomes (1, 17, 18). Our findings raise the possibility that at least, in some settings, and for some changes, maintaining a high-level of change commitment may be immaterial for generating implementation effort. Our findings also raise questions about change efficacy, as we failed to find change efficacy associated with implementation effort. That may be due to issues with construct validity; unreliable measurement; sampling bias; or a combination. However, given our careful survey development and validation procedure (7) and the rigorous experimental design of these findings, at the very least, these results place a burden of proof on future studies in workplace health promotion that rely on change efficacy as a mechanism for change to demonstrate construct validity and measurement reliability.

Third and finally, many of our current implementation models and measures focus on attitudinal constructs, such as commitment, efficacy, and motivation, but this study suggests that more instrumental constructs, such as the planning and technical support that was provided by HealthLinks, may be more important variables in ensuring effective implementation. As noted, behavioral economics has repeatedly shown that people are often poor at predicting their own behaviors, or acting in ways that are consistent with their expressed goals and self-interest (19, 20). This experiment needs to be replicated, but if our findings are reproduced, one implication may be that measuring and influencing affective states is less useful than ensuring instrumental support, such as planning, which runs counter to some of the current thinking in the literature (21).

**Limitations**

This study has several limitations that raise a variety of interesting questions. First, we found change commitment declined across study arms, likely due to regression to the mean. Our findings about change commitment also might reflect selection bias. The study population by definition only included volunteers, who were virtually certain to exhibit higher-than-average change commitment. This may have constrained the observed variation in change commitment. If we were able to randomize the whole population of small worksites in low-wage industries in King County to HealthLinks or control conditions, it is possible that we would observe significant changes in change commitment over time and significant differences between HealthLinks and control sites.

Second, baseline readiness factors notably change commitment and change efficacy might still be significant predictors of subsequent wellness program effort irrespective of the plasticity of the measures over time or their sensitivity to the effects of implementation strategies, such as HealthLinks. For example, baseline readiness factors including change commitment and change efficacy might be important independent predictors of subsequent wellness program effort. Or, they might be necessary but not sufficient conditions for successful implementation, and we could observe significant interactions between readiness factors and implementation strategy, such that sites with a high baseline-level of change commitment and change efficacy AND who receive the HealthLinks intervention would demonstrate much higher levels of wellness program effort than either sites with high baseline-level of change commitment and change efficacy OR receipt of the HealthLinks intervention alone. These questions were beyond the scope of the current analysis and are the focus of future work.

It is also possible that we need to rethink our conceptualization of readiness to change. At baseline, respondents were rating hypotheticals: how committed were they to engaging in a set of practices with which they generally did not have prior experience? How confident were they in their collective ability to implement health promotion practices? Research in cognitive psychology and behavioral economics has repeatedly shown people to be poor at predicting future behaviors, states, and feelings (19, 20). Participant ratings of their readiness to implement a new practice might be inherently unreliable until they have gained some experience with the practice. An alternative approach that could be tested in the future is to have participants estimate base rates: when they or others in their industry have attempted similar initiatives in the past, how often were they successful, and what were the main stumbling blocks and facilitators?

Our findings may have been biased by measurement error. The survey was fielded to a single individual, typically a human resources manager, identified by the employer as the contact for the study. Weiner’s theory (1) postulated that readiness is a shared construct, and ideally would be measured among all employees involved in the change. It is possible that the individuals in our sample had incomplete or flawed insights into their companies’ readiness domains, and that a different sample, e.g., a broader sample of employees, or company executives, would produce more accurate measures of readiness and a different result. In more than a third of participating sites, there was turnover in the primary study contact completing these measures, and this could also introduce measurement error. Important questions for future research are to what degree there is agreement among employees within workplaces about the level of change commitment and change efficacy, and whether level of agreement itself may be a predictor of implementation.

Finally, the intervention (HealthLinks), the target practice (workplace health promotion practices), and setting (worksites...
in low-wage industries in King County, Washington State) may limit the generalizability of the findings. However, there are not theoretical reasons we are aware of that would explain why change commitment and change efficacy would be unrelated to change effort in this context but should be in other contexts.

While our study had limitations, it also had important strengths. We do not know of other studies that have (1) systematically developed and independently validated (including item comprehension, construct validity, scale reliability, and criterion validity) a theory-based measure tailor-made for a specific implementation program and setting; (2) prospectively assessed changes in readiness with measures of program-change effort; and (3) used experimentally manipulated conditions directed at changing readiness factors. We believe this design made for a unique, scientifically rigorous study.

Ultimately, these findings raise more questions than they answer and point to a number of interesting avenues for future research.

CONCLUSION

Many implementation theories predict that commitment and efficacy mediate the effect of implementation strategies and actual implementation efforts. We did not find this to be the case in the setting of small worksites in low-wage industries implementing evidence-based health promotion practices. Instead, we found implementation strategies can lead to significant implementation efforts in the absence of improved change commitment—indeed, in the presence of declining change commitment. If replicated—at least in this setting—this suggests that implementation measures and models may be better served by focusing less on attitudinal constructs and more on instrumental constructs, such as planning and technical support.

AUTHOR CONTRIBUTIONS

CH, MK, CA, KH, KC, AP, BW, JH, and PH contributed substantially to the conception and design of the study and/or to the acquisition of data. All authors contributed substantially to the analysis and interpretation of findings. CH and PH drafted the article; all authors provided critical revision of the article. All authors provided final approval of the version to be published and all agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

We would like to acknowledge the HealthLinks employers and employees for participating; and our collaboration with the American Cancer Society. We are grateful to our partners and funder.

FUNDING

This study was funded by grant 5R01CA160217 from the National Cancer Institute.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fpubh.2018.00110/full#supplementary-material.

REFERENCES


**Disclaimer:** The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs, the National Cancer Institute, or the United States Government.

**Conflict of Interest Statement:** The authors declare the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer DH declared a shared affiliation, with no collaboration, with the authors to the handling Editor.

Copyright © 2018 Helfrich, Kohn, Stapleton, Allen, Hammerback, Chan, Parrish, Ryan, Weiner, Harris and Hannon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Systematic, Multimethod Assessment of Adaptations Across Four Diverse Health Systems Interventions

Borsika A. Rabin1,2,3,4*, Marina McCreight1, Catherine Battaglia1,5, Roman Ayele1,5, Robert E. Burke1,6, Paul L. Hess1,6, Joseph W. Frank1,6 and Russell E. Glasgow1,3,4

1 Denver-Seattle Center of Innovation for Veteran-Centered and Value-Driven Care (COIN), Denver VHA Medical Center, Denver, CO, United States, 2 Department of Family Medicine and Public Health, School of Medicine, University of California San Diego, La Jolla, CA, United States, 3 Adult and Child Consortium for Health Outcomes Research and Delivery Science, School of Medicine, University of Colorado, Aurora, CO, United States, 4 Department of Family Medicine, School of Medicine, University of Colorado, Aurora, CO, United States, 5 Department of Health System Management and Policy, Colorado School of Public Health, University of Colorado, Aurora, CO, United States, 6 Department of Medicine, School of Medicine, University of Colorado, Aurora, CO, United States

Background: Many health outcomes and implementation science studies have demonstrated the importance of tailoring evidence-based care interventions to local context to improve fit. By adapting to local culture, history, resources, characteristics, and priorities, interventions are more likely to lead to improved outcomes. However, it is unclear how best to adapt evidence-based programs and promising innovations. There are few guides or examples of how to best categorize or assess health-care adaptations, and even fewer that are brief and practical for use by non-researchers.

Materials and methods: This study describes the importance and potential of assessing adaptations before, during, and after the implementation of health systems interventions. We present a promising multilevel and multimethod approach developed and being applied across four different health systems interventions. Finally, we discuss implications and opportunities for future research.

Results: The four case studies are diverse in the conditions addressed, interventions, and implementation strategies. They include two nurse coordinator-based transition of care interventions, a data and training-driven multimodal pain management project, and a cardiovascular patient-reported outcomes project, all of which are using audit and feedback. We used the same modified adaptation framework to document changes made to the interventions and implementation strategies. To create the modified framework, we started with the adaptation and modification model developed by Stirman and colleagues and expanded it by adding concepts from the RE-AIM framework. Our assessments address the intuitive domains of Who, How, When, What, and Why to classify and organize adaptations. For each case study, we discuss how the modified framework was operationalized, the multiple methods used to collect data, results to date and approaches utilized for data analysis. These methods include a real-time tracking system and structured interviews at key times during the intervention. We provide...
INTRODUCTION
Implementing a program is like constructing a building. An architect draws upon general engineering principles (theory) to design a building that will serve the purposes for which it is designed. However, the specific building that results is strongly influenced by parameters of the building site, such as the lot size, the nature of the site’s geological features, the composition of the soil, the incline of the surface, the stability and extremes of climate, zoning regulations, and cost of labor and materials. The architect must combine architectural principles with site parameters to design a specific building for a specific purpose on a specific site. This dynamic is mirrored in the rough-and-tumble world of the human services. Despite excellent plans and experience, ongoing redesign and adjustment may be necessary. [Bauman et al., (1)]

Health systems interventions are rarely ever implemented in precisely the same way across diverse, real-world settings. Changes to the original intervention and/or implementation protocol during the course of a program are described as adaptations in the dissemination and implementation literature and are receiving growing attention from researchers and practitioners alike (2, 3). By considering local culture, history, resources, characteristics, and priorities, interventions are more likely to lead to improved outcomes (2-5). Understanding the nature, origin, timing, and impact of these adaptations is crucial for many reasons. Adaptation information can provide contextual and process data and support the interpretation of study findings. It can also help identify which components of the intervention and implementation strategies worked and which components need to be modified in a given setting and for a given population, and can ultimately help answer the question of what components of an intervention work for what population, for producing what outcomes, under what circumstances. The information can then guide real-time or end-of-project improvements and refinements to intervention and implementation strategies and provides guidance for future scale up and scale out (6).

A critical piece in identifying adaptations to an intervention and implementation protocol is to find strategies to systematically evaluate and document the adaptations. The ideal pragmatic approach to documenting and evaluating adaptations happens in real time and throughout the lifetime of the project, is replicable, is unobtrusive to the users and beneficiaries of the intervention, has low complexity, is low cost and requires modest resources, provides both quantitative and qualitative information on the adaptation, assesses the adaptations from the perspective of multiple stakeholders, and uses multiple methods to generate rich data (7). Furthermore, an assessment strategy that can be applied across diverse settings, interventions, and implementation strategies would permit and encourage cross-study comparisons. Finding such an approach or combination of approaches poses a challenge, and there is little guidance in the literature to date.

Given the novelty of the field of adaptation research, there are numerous opportunities to develop and test methods to address questions such as which types of adaptations are most beneficial and which result in reduced fidelity and worse outcomes (2, 3). For example, are adaptations made before implementation any more or less helpful; are intentional adaptations more productive than unintentional ones; and are externally required (versus internally motivated) adaptations more disruptive?

In this study, we describe a mixed and multimethod approach to documenting and evaluating adaptations in the context of four, diverse, multisite health systems interventions and implementation efforts that are being applied in the Veterans Health Administration (VHA) health-care system. We describe our adaptation documentation and evaluation strategies, including a modified framework and multiple methods used to collect data, provide preliminary findings on adaptations from four health systems intervention and implementation studies, and share lessons learned and possible applications of our methodology. Our assessment methods below are in the public domain and are available upon request from the authors, and we encourage their use, evaluation, and improvement.

METHODS
Section “Methods” provides a description of our four interventions, implementation strategies, and their settings; the adaptation framework and coding system used to guide our documentation and evaluation activities, and the details of the documentation and evaluation approach used across the four case studies.

Setting and the Four Interventions
The VHA is the largest integrated health care system in the United States, providing primary and specialty health services to nine million enrolled Veterans. The VHA plays a lead role in

descriptive data on the types and categories of adaptations made and discuss lessons learned.

Conclusion: The multimethod approaches demonstrate utility across diverse health systems interventions. The modified adaptations model adequately captures adaptations across the various projects and content areas. We recommend systematic documentation of adaptations in future clinical and public health research and have made our assessment materials publicly available.

Keywords: adaptation, RE-AIM framework, Stirman framework, mixed methods, pragmatic measures, assessment
improving the quality of patient care and health services through multiple initiatives, and the Quality Enhancement Research Initiative (QUERI) has been a central component of the VHA’s commitment to improve health care for Veterans (8). The Triple Aim QUERI is 1 of 15 currently funded QUERI programs and focuses on leveraging health-care data to identify actionable gaps in care, and to implement innovative health-care delivery interventions to improve the Triple Aims of VHA health care which are patient-centered care, population health, and value. The Triple Aim QUERI uses three projects to assess the feasibility and effectiveness of various interventions and implementation strategies unified by shared implementation models, measures, and approaches. In addition to these three projects, this manuscript includes a project from a sister VHA initiative funded through the VHA Office of Rural Health.

The four projects are described with their key characteristics in Table 1. As shown in Table 1, the four projects are diverse in the program focus area, clinical problem they address, target population, and the intervention format and delivery. The first project, titled Implementation of Extensible Methods to Capture, Report, and Improve Patient Health Status (Patient Reported Health Status Assessment), aims to utilize and implement the interactive voice response (IVR) to capture the pre- and postprocedural patient-reported health status for patients receiving elective catheterization laboratory procedures with intent to inform clinical care (9). The second project, titled Leveraging Data to Improve Multimodal Pain Care Through Targeted Telementoring (Multimodal Pain), aims to address barrier and facilitators to multimodal pain care in the VHA and to design and implement an intervention based on identified best practices to support primary care providers (10). The third project, titled Improving Veterans Transition Back to VA Primary Care Following Non-VHA Hospitalization (Community Transitions), focuses on care coordination of those Veterans admitted to non-VHA community hospitals for inpatient care and transition back to VHA primary care in a safe, patient-centered and timely manner (11). The fourth project, the Transitions Nurse Program (Rural Transitions), is a proactive, personalized, nurse-led and Veteran-centered intervention to improve access for rural Veterans to follow-up with their PACT teams following hospitalization at a larger urban VHA Medical Center (VAMC) (12).
These four projects involved diverse groups of local, regional, and national operational partners from the inception of the projects. As part of this effort, each project actively engaged key operational partners and identified outcomes of direct relevance to these partners. The Multimodal Pain project partnered with the Office of Specialty Care and National Program for Pain Management; the Patient Reported Health Assessment project teamed with National Cardiology Program, Clinical Assessment Reporting and Tracking Program, Office of Analytics and Business Intelligence, Office of Quality, Safety and Value; the Community Transition project teamed with VHA Office of Community Care and VISN 19 Rural Health Resource Center-Western Region; and the Rural Transition project partner with the Office of Rural Health and the Office of Nursing Services. Furthermore, each program utilized the Denver VHA Veteran Research Engagement Board. The Engagement Board brings Veterans and other health-care system stakeholders together to contribute to research in meaningful ways.

We involved Veterans at multiple phases of the project, including the design, implementation, adaptation, and evaluation. Individual projects have the opportunity to speak to Veterans from diverse socioeconomic and service backgrounds and receive rapid feedback and questioning to ensure the program being implemented has positive impact on Veterans, providers, and their care givers.

We also involved local VHA and non-VHA stakeholders where we learned about barriers and facilitators to current processes at the VHA and obtained suggestions for improvement. For example, the Community Transitions project teams conducted in-depth, pre-implementation assessment of the current process with VHA and non-VHA clinicians and staff as well as Veterans to understand the current transition of care process. Following this assessment, an intervention was designed to address barriers identified by these VHA and non-VHA participants. During the implementation phase, project team members reached out to VHA and community stakeholders to describe the intervention, its value to those involved and answer questions. During these meetings, project sub-teams were asked to tweak certain elements of the intervention that they then brought back to the larger team to discuss feasibility, value added and if it would improve health outcomes for Veterans. This iterative process continues as the intervention is ongoing and new community stakeholders are engaged. To keep adaptation information organized, each interaction is documented including the source of information, date of suggested change or improvement and their care givers.

We involved local VHA and non-VHA stakeholders where we learned about barriers and facilitators to current processes at the VHA and obtained suggestions for improvement. For example, the Community Transitions project teams conducted in-depth, pre-implementation assessment of the current process with VHA and non-VHA clinicians and staff as well as Veterans to understand the current transition of care process. Following this assessment, an intervention was designed to address barriers identified by these VHA and non-VHA participants. During the implementation phase, project team members reached out to VHA and community stakeholders to describe the intervention, its value to those involved and answer questions. During these meetings, project sub-teams were asked to tweak certain elements of the intervention that they then brought back to the larger team to discuss feasibility, value added and if it would improve health outcomes for Veterans. This iterative process continues as the intervention is ongoing and new community stakeholders are engaged. To keep adaptation information organized, each interaction is documented including the source of information, date of suggested change or improvement and comments.

This study was not considered research per VHA ORO policy 1058.05, therefore ethical review and approval was not required in accordance with the local legislation and institutional guidelines.

**Adaptation Framework**

A number of adaptation frameworks currently exist. Many of them originated from the cultural adaptations literature that first acknowledged that interventions needed to be appropriately adapted to fit local cultural needs to be successful (2, 5). A systematic effort was conducted by Stirman and colleagues to identify the core characteristics of adaptations and modifications in the dissemination and implementation literature and resulted in the coding guide we reference as the Stirman adaptation and modification framework (4). The original Stirman framework provides a method to systematically code adaptations made to the content of the intervention (nature and level) and to the context in which the intervention is delivered as well as to document by whom the adaptations were made (4). Hall and colleagues affiliated with our research group investigated adaptations in the primary care setting and found that to fully capture the nature and impact of adaptations in those applied settings it was necessary to expand the Stirman et al framework (13). They found the original Stirman framework categories useful, but further expanded the framework by adding constructs informed by the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework\(^2\) to include why and when the adaptations were made and what the impact of the adaptations were (13, 14). The core constructs of the modified adaptation framework are described in Table 2. For ease of use and understanding by clinical and community leaders and staff who were interviewed, these domains were framed using intuitive categories of Who, How, When, What, and Why to classify and organize adaptations. For each area, coding categories are identified and listed in the table. This framework and coding system is used to inform the documentation and evaluation approach described in the next section.

**Documentation and Evaluation Approach**

Our documentation and evaluation approach has two main components. We first created a robust documentation tool allowing for the real-time, ongoing tracking of adaptations throughout the course of the project, and we also used a semi-structured, multilevel, and multistakeholder interviews implemented at multiple time points. The combination of these two approaches is intended to provide rich data on adaptations to the intervention and implementation strategies, and inform the subsequent expansion of the intervention to additional sites in the VHA. Each of these approaches is described below in more detail. Lessons learned from the implementation of these approaches to date are summarized in Section “Results.”

**Real-Time and Ongoing Tracking of Adaptations**

The adapted Stirman framework and coding system was used to create a pragmatic, easy-to-use tabular worksheet to track adaptations as they occurred throughout the lifetime of the project. The original worksheet was pilot tested and refined to improve usability and decrease burden and obtrusiveness. The current version of the worksheet is used by project research personnel (i.e., project manager or coordinator) and is presented in Table 3 along with two examples of recorded adaptations. The real-time tracking sheet is designed to be used from the early planning stages of the project and is populated on a regular basis in consultation with frontline implementers. The goal of this assessment method is to allow for comprehensive capturing of changes made.

---

Table 2: The Triple Aim Quality Enhancement Research Initiative Adapted Stirman Adaptation framework and coding system and interview questions.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Coding categories</th>
<th>Interview question example</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHAT is modified?</td>
<td>Content (modifications made to content itself, or that impact how aspects of the treatment are delivered)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Context (modifications made to the way the overall treatment is delivered)</td>
<td>WHAT Part 1: WHAT component or part of the intervention was changed in this adaptation; in other words, what was the nature of the change? For instance, was it a change to program content, format, delivery mode, staff delivering it, patients eligible, where, when or how it was delivered, or what?</td>
</tr>
<tr>
<td></td>
<td>Training and evaluation (modifications made to the way that staff are trained in or how the intervention is evaluated)</td>
<td></td>
</tr>
</tbody>
</table>
| At what LEVEL OF DELIVERY (for whom/what are modifications made?) | – Individual patient level  
– Group level  
– Individual practitioner level  
– Clinic/unit level  
– Hospital level  
– Network level  
– System Level | Implied from other questions                                                                 |
| Context modifications are made to which of the following? | – Format  
– Setting  
– Personnel  
– Population | Was the change to program, content, format, delivery mode, staff delivering it, patients eligible, where, when, or how it was delivered or what? |
| What is the nature of the content modification? | – Tailoring/tweaking/refining  
– Adding elements  
– Removing/skipping elements  
– Shortening/condensing (pacing/timing)  
– Lengthening/extending (pacing/timing)  
– Substituting  
– Reordering of intervention modules or segments  
– Integrating the intervention into another framework (e.g., selecting elements)  
– Integrating another treatment into EBP (not using the whole protocol and integrating other techniques into a general EBP approach)  
– Repeating elements or modules  
– Loosening structure  
– Departing from the intervention (“drift”) | WHAT Part 2: How would you describe the type of change involved in this adaptation? Specifically, what did the change involve? Was something added, deleted, changed to better fit the patients, delivered at a different time or in a different way? |
| WHEN: When during the project the adaptation was made? | – During planning stages, before intervention began  
– Early, during first few weeks of intervention  
– During the middle stages  
– At or close to the end of project | WHEN during the ____ program was this adaptation first made? |
| HOW: How or on what BASIS was this change made? | – Based on our vision or values  
– Based on a framework (for example, PCMH)  
– Based on our knowledge or experience of working with patients  
– Based on QI data, summary information or results  
– Based on pragmatic/practical considerations (for example, “this is the only way it would work”)  
– Based on financial incentives/payment  
– Based on feedback or suggestions (Practice Facilitator/coach or other)  
– Other | HOW or on what BASIS was this change made—based on challenges implementing, on time concerns, on results or data you collected, on external or administrative concerns, feedback from patients or staff, or what basis? |
| WHY: What was the purpose of the adaptation? | – Increase reach, participation, access  
– Increase effectiveness  
– Increase adoption by more clinics/settings or make intervention more aligned with organizational goals  
– Increase implementation/ability of staff to deliver intervention successfully | 1. WHY Part 1: WHY was this adaptation made? For example, to get more people to participate, to make the program attractive to more settings, to increase its effectiveness, to make it easier to deliver, to make it easier to maintain or reduce costs, etc.?  
2. WHY Part 2: Was this adaptation a result of EXTERNAL factors (for example, change in organizational policies, reimbursement changes) or INTERNAL issues, such as workflow, changes in staff or similar issues? |
| BY WHOM are modifications made? | – Individual practitioner/facilitator  
– Team  
– Non-program staff  
– Administration  
– Program developer/purveyor  
– Researcher  
– Coalition of stakeholders  
– Unknown/unspecified | WHO was responsible for first suggesting or initiating this change? Was this the person or persons the ones who implemented the change? (If not, who implemented the adaptation?) |

(Continued)
to the project and to improve recall during adaptation interviews described below.

The implementation teams used different strategies to support the implementation of the real-time tracking form across our four projects. These strategies included first, the addition of a standing agenda item to weekly/biweekly meetings with implementers to ask about challenges they encountered during the implementation of the project and whether they needed to make or planning on making any changes to address these challenges; the discussion and adaptations data collection was facilitated by both implementation and clinical team leads. Second, some projects converted their regular team meeting documents (such as action items and minutes) into data that fit into the main delivery were used to clarify, add to, or enhance adaptation descriptions; these included the field notes and process maps from site visits. In this project, a team consisting of an implementation specialist and a research nurse conducted site visits to all expansion sites approximately 6 months after intervention initiation to directly observe the delivery of the intervention and document adaptations made since program roll-out at each site. The observational data are used to construct intervention process maps and provide additional contextual factors for the implementation evaluation. The remaining projects are planning to adopt this approach when the interventions are expanded to additional sites. Information from the real-time tracking system is used to create a list of adaptations as well as to support interviews (i.e., help with recall).

**Semi-Structured, Multilevel, and Multistakeholder Interviews**

A semi-structured interview guide and coding system adapted from that used by Hall and colleagues (13) tailored to the context of our four projects was drafted and pilot tested. Example questions and probes from the interview guide as they align with the various construct/coding categories are listed in Table 2. First, interviewees are asked to identify all changes they made to the original intervention or implementation strategy protocol. Then, they are asked to identify the most important changes made to the intervention or implementation strategy and to list them in the order of perceived importance, with the first change being most important. Detailed follow-up questions are then asked related to the change that was deemed most important by the interviewee; if

### TABLE 2 | Continued

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Coding categories</th>
<th>Interview question example</th>
</tr>
</thead>
<tbody>
<tr>
<td>*IMPACT: What are (subjective) short-term results of the adaptation?</td>
<td>- Are they positive, negative, no real impact? - Did the changes impact: o Reach/participation/access o Effectiveness o Adoption o Implementation/ability of staff to deliver intervention successfully o Maintenance</td>
<td>What was the short-term IMPACT of this adaptation? Did it have highly visible results? (For example did it result in more or less participation by patients, get more or fewer settings or staff involved, improve or decrease consistency of delivery, improve or reduce outcomes, reduce or increase time or costs? We understand that you may not have concrete outcomes results at this time—please tell us your best perception of the impact of this adaptation thus far.)</td>
</tr>
</tbody>
</table>

*Additional constructs to original Stirman Adaptation Framework.

### TABLE 3 | Real-time tracking of adaptations form and two examples.

<table>
<thead>
<tr>
<th>Date of the modification</th>
<th>4/15/2016</th>
<th>6/2/2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description of the specific modification</td>
<td>ISurvey questions reordered—moved the Rose Dyspnea questionnaire to the end</td>
<td>Revised patient letter to include information about automated pre-procedural phone calls</td>
</tr>
<tr>
<td>Reason for the modification</td>
<td>To improve fluidity of the survey and enhance data capture</td>
<td>To prepare patients for data collection</td>
</tr>
<tr>
<td>BY WHOM are modifications made?</td>
<td>Researcher</td>
<td>Researcher</td>
</tr>
<tr>
<td>WHAT is modified?</td>
<td>Order of data collection</td>
<td>Content of the intervention</td>
</tr>
<tr>
<td>At what LEVEL OF DELIVERY?</td>
<td>Individual patient level</td>
<td>Individual patient level</td>
</tr>
<tr>
<td>CONTEXT modifications are made to...</td>
<td>Intervention format</td>
<td>Intervention format</td>
</tr>
<tr>
<td>What is the NATURE of the Content modification?</td>
<td>Tailoring/tweaking/refining</td>
<td>Tailoring/tweaking/refining</td>
</tr>
<tr>
<td>WHEN: When during the project the adaptation was made</td>
<td>During planning stages before began intervention</td>
<td>During planning stages before began intervention</td>
</tr>
<tr>
<td>WHY: What is the purpose of the adaptation?</td>
<td>Increase effectiveness</td>
<td>Increase implementation/ability of staff to deliver intervention successfully</td>
</tr>
<tr>
<td>IMPACT—What are (subjective) short-term results of adaptation?</td>
<td>Positive: impact effectiveness</td>
<td>Positive: impact implementation/ability of staff to deliver intervention successfully</td>
</tr>
</tbody>
</table>
time permits, follow-up questions are asked about the additional changes mentioned in the beginning of the interview. In some cases, adaptations documented in the real-time tracking document were systematically used to improve interviewee recall and remind interviewees about important changes that happened during the implementation of the intervention. The semi-structured adaptation interviews are designed to be conducted at two time points or more during the project, including soon after implementation of the intervention (within 3–6 months) and at the end of the project. The full interview is available at https://goo.gl/PDGWtf.

Each project identifies a set of stakeholders to interview including frontline implementers and research personnel. Interviews are audio recorded, transcribed, and coded. The qualitative content is managed using Atlas ti. software package. The qualitative analytical team uses consensus-building to discuss the emergent codes and themes and to resolve differences in coding. Data are summarized in the form of adaptation lists. Each project plans to conduct two waves of interviews, one soon after implementation and another right at the end of the project. We are planning on interviewing up to 10 people in various roles in the implementation process for each wave and project. Findings from the earlier wave of the interviews will be used to inform refinements to our interventions and implementation strategies and approaches for subsequent expansion of the interventions as well as to support interpretation of our findings at the end of the project. We will also use information from these interviews (in combination with the data emerging from the real-time tracking system) to create an adaptation guide for future implementers.

RESULTS

In this section, we share preliminary results and lessons learned from our four projects. All four of these projects are in progress and at various stages of the planning and implementation continuum.

Real-Time and Ongoing Tracking of Adaptations

The real-time tracking system has been implemented across all four projects. We have documented a total of 46 adaptations to date across the four projects (average of 12 per project, most of which occurred shortly after initiation of the intervention). Table 3 lists two specific examples and demonstrates what the real-time tracking document look like in action. Most adaptations documented to date have been related to the intervention delivery, such as defining and fine-tuning enrollment criteria in the Rural Transitions project, initiation of the IVR calls in the Patient Reported Health Assessment project, and recruitment materials in the Community Transitions project.

The real-time tracking sheet is used by project managers or coordinators on a weekly basis. It requires approximately 3–5 min to complete the tracking sheet for each adaptation. Key adaptations documented here included scope of the intervention, its delivery and evaluation plans for the Community Transitions project, expansion of the enrollment criteria in the Rural Transitions project; modifications to the IVR calls delivery in the Patient Reported Health Assessment project. Some key lessons learned from the use of the tracking sheet are summarized in Table 4. Positive feedback included the perceived usefulness of documenting information in a structured manner which allows for ready retrieval at a later time, and help with identifying core components of the intervention and implementation protocol.

Some of the lessons include strategies on how to implement the real-time tracking system [e.g., the need to set reminders (calendar reminder)], the importance of checking in regularly with the project team along with using the tracking sheet, and the need to communicate with frontline implementers about possible changes/adaptations as the research team is not always aware of all changes made by frontline staff. Another challenge that was identified was that results of adaptations may not be clear until weeks or months after the change, making it difficult to record this information.

Semi-structured, Multilevel, and Multistakeholder Interviews

Adaptation interviews have started in three of the four projects (Rural Transitions, Community Transitions, and Patient Reported Health Assessment). We conducted 11 interviews with site implementers (transitions nurses and champions) in Rural Transitions, three interviews with program staff and transitions nurse in Community Transitions; and one interview with implementers in the Patient Reported Health Assessment project. Interviews last an average of 45 min. Table 4 summarizes early lessons from our adaptation interviews. Key reflections include the realization that some interview questions might not be different enough to produce distinct responses (e.g., questions about WHY and HOW), the sequence of the interview was not always optimal (e.g., would prefer to ask details about adaptation when first mentioned instead of waiting to list all adaptations), probes were helpful in most cases, the introduction for the interview was too lengthy, and it can be challenging to record information about the adaptation in the interview table while conducting the interview. One unexpected finding from these early interviews was that an adaptation of the Rural Transitions intervention was to limit the number of eligible patients to enroll to avoid the burnout of the transitions nurse. Information from these early interviews has been used to inform the intervention roll-out process for the subsequent expansion of the Rural Transitions program by providing guidance on the enrollment strategies for the on-coming sites. Finally, the timing for conducting the early wave of interviews were somewhat delayed by the competing demands of the implementation of the intervention.

DISCUSSION

Our adaptations project has conceptualized assessment methods, developed and adapted multimethod procedures, and is applying them across four diverse projects and content areas. The methods appear to be feasible, informative, and applicable across different clinical targets, interventions, research projects, and settings. As discussed below, preliminary results appear to be promising.
and investigations are ongoing. Our focus throughout, in accord with implementation and dissemination principles (15, 16), has been on multiple methods, multiple contextual levels, and rapid, pragmatic assessment strategies (17). We summarize overall experiences to date, lessons learned, strengths and limitations of pragmatic assessment, which has recently received increased attention in implementation science (7, 18–23). Importantly, busy clinical staff are not asked to complete lengthy questionnaires or spend lots of time in added meetings or assessment procedures. The most time-consuming activities including tracking records, conducting and analyzing interviews, and conducting observations can be completed by project managers or research assistants without high levels of advanced education. Many activities, especially the tracking documentation, can be accomplished by keeping good records during existing project management and supervision activities.

Our assessment methods are flexible and can be tailored to different projects and purposes. They can be adapted to a particular project in terms of the sources and levels of information collected (e.g., CEOs and macro-level adoption decisions; providers and guidelines application; front line delivery staff and implementation actions). Tracking and observational data can be collected in the context of any combination of team meetings, site visits, other assessment procedures, quality control contacts, direct observations, phone call check-ins or other opportunities. These rapid and frequent assessment methods can be used iteratively to inform future inquiries and adaptations. Thus far, we have made use of this feature by tracking data to be assessed in more detail in structured adaptation interviews (13).

### TABLE 4 | Lessons learned from using the real-time tracking system and interviews.

<table>
<thead>
<tr>
<th>Project</th>
<th>Project role</th>
<th>Tracking tool used</th>
<th>Reflections/feedback/lessons learned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Transitions</td>
<td>Project coordinator</td>
<td>Real-time tracking form</td>
<td>Easy to forget real-time tracking, recurring reminders are useful</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Helps when need to go back and look at certain decisions made with the team</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Easy to forget if there are a lot of changes happening to the intervention in a short period of time</td>
</tr>
<tr>
<td>Patient Reported Health Status Assessment</td>
<td>Project coordinator</td>
<td>Real-time tracking form</td>
<td>Helps clarify core intervention and implementation components</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Critical to talk to line implementers—decision-maker and research team may be unaware of adaptations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Helps document exactly what component was adapted—when, why, how, and by whom (Stirman framework as a guide)</td>
</tr>
<tr>
<td>Multimodal Pain</td>
<td>Project coordinator</td>
<td>Real-time tracking form</td>
<td>Helps keep audit trails up to date so time is not wasted digging through emails/notes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Need to organize inbox folders by project and keep emails related to the project</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Need to set reminders in outlook (or other system) to update tracking sheets</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Important to document everything even if it does not seem important at the time</td>
</tr>
<tr>
<td>Rural Transitions</td>
<td>Implementation specialist</td>
<td>Real-time tracking form</td>
<td>Challenging to track in real-time simultaneously at multiple sites and based on feedback from multiple team members</td>
</tr>
<tr>
<td>Rural Transitions</td>
<td>Research Transitions coordinator, RN</td>
<td>Real-time tracking form</td>
<td>Easy to check in with teams as it became a standing part of the agenda</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- New way to track changes in the project—was not familiar and as convenient</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Useful system and provided valuable information to refer back to</td>
</tr>
<tr>
<td>Rural Transitions</td>
<td>Qualitative analyst 3</td>
<td>Real-time database tracking</td>
<td>Need to remind the site implementation teams to track in the database</td>
</tr>
<tr>
<td>Rural Transitions</td>
<td>Qualitative analyst 1</td>
<td>Adaptation interview guide</td>
<td>Introduction to the interview was too long</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Difficult to have interviewees first list all adaptations and then go through the questions for each adaptation. Would be more organic to have them mention the first adaptation, then ask follow-up questions, then probe for more adaptations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Some of the probing questions felt repetitive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Might be helpful to do interviews in the first couple months of implementation since that is when most of the adaptations seemed to take place</td>
</tr>
<tr>
<td>Rural Transitions</td>
<td>Qualitative analyst 2</td>
<td>Adaptation interview guide</td>
<td>Interview was too long</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Participant did not feel they had made adaptations, so some of the follow-up questions were awkward</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Difficult to fill in table while conducting interview, need to code after interview is complete</td>
</tr>
<tr>
<td>Rural Transitions</td>
<td>Qualitative analyst 3</td>
<td>Adaptation interview guide</td>
<td>Interview questions needed additional clarifications, so the follow-up questions and probes were helpful to clarify</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Some questions seemed repetitive (e.g., How and Why)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Many times, an interviewee would start describing the details of the adaptations and answer all the follow-up questions without prompting</td>
</tr>
</tbody>
</table>
An optional feature of our assessment methods can be viewed as either a strength or limitation. On the one hand, the procedures, coding categories, areas of focus, and results assessed can vary over time and are informed by accumulating data. From a traditional efficacy research and psychometric perspective, some of these updates and assessment modifications may be seen as methodologically problematic. From this perspective, assessment should be defined before data collection and applied in a standardized fashion regardless of results (and results may not be reviewed until project conclusion). We understand this perspective, and note that flexible and iterative use of our methods is optional and not required if these features are not desired. On the other hand, in the spirit of rapid use of research results and improvement science (24), actionable information can and should generally be useful to inform intervention and implementation adaptations, which are likely to occur in any case, but otherwise be less informed by data (3).

This study and our methods raise two important additional questions (1) why, how, what types of adaptations are successful (or not) and (2) how might one “optimize” adaptation of an intervention at the design stage for maximum success. We do not yet have data on the first issue but will at the conclusion of the four projects. We do collect “immediate perceived impact” of the staff and interviewee on the forms, but these are subjective and do not address delayed effects. The second issue of the use of these adaptation data on how to optimize intervention—and implementation strategies—is clearly important and extremely complex (25). Some researchers do not feel it is appropriate to modify an intervention following development of an initial protocol, and others have proposed both adaptive or SMART designs and use of modeling approaches to address these issues (26). More detailed discussion of these issues is provided, for example, in Riley and Rivera (27).

Our experiences to date have also revealed challenges and limitations to these assessment procedures. First, optimal use of our multiple methods requires in-depth knowledge of project intervention and implementation strategies. Sometimes assessment staff are not in contact with intervention planners or implementers and are not informed about procedural details. Our methods can still be used in such situations, but will likely not be as specifically useful to those projects in terms of informing future directions. Our team had initial difficulties in differentiating adaptations made to intervention components versus implementation strategies (such as audit and feedback or facilitation) (28), partially because the grant project applications funding our assessments were not always clear on these distinctions. Our methods can be used to assess adaptation to either or both intervention components or implementation strategies. In some cases, these distinctions may be important for either scientific or application purposes; in other cases, they may not. Furthermore, our current project only allowed for the administration of the interview portion of our methodology at two time points. An additional interview during the planning/pre-implementation phase would be ideal but not essential.

Since the projects involved had moderately specific protocols concerning intervention components, and especially implementation strategies (rather than being scripted and manualized interventions), it was sometime challenging to understand precisely what the intervention component was and whether it was adapted or implemented as originally intended. Other limitations include that thus far we have not conducted formal reliability or validity assessments.

Our adaptation assessment methods are based upon the Sitraman and RE-AIM frameworks, both of which have been used in multiple settings and found valid and useful (4, 13, 29, 30). However, the specific assessment instruments used in this study, while demonstrating high face validity, have not been subjected to formal psychometric testing. Since these are new measures, the analytic implications of these methods are unclear. Many potentially useful variables (e.g., timing, source, content, and purpose of adaptation) can be coded from these methods, but it is not clear which are most important, their interrelationships, or exactly how they should be analyzed (e.g., continuous versus dichotomous variables).

Furthermore, it does take time and effort to collect these adaptation data and their value needs to be weighed against alternative uses of resources. We have tried to minimize the time and burden on both staff (e.g., reccoding tracking form data during regular meetings) and delivery staff (doing only two interviews at convenient times), but these activities might not be high enough priorities for some projects to justify the time. Another important consideration is the way in which impact is tracked across time using the proposed approach. While we do not systematically follow-up on tracking data to evaluate the impact of adaptations (we do assess initial impact, but some adaptations are of course delayed in time), there are concurrent assessments of some separate process and intermediate outcome measures.

Finally, as is the case with multiple methods in general, it is not clear exactly how to integrate data from multiple sources (31–33). Despite these limitations, we conclude that these multiple adaptation assessment methods are useful and worthy of further investigation. In addition to formal psychometric testing regarding reliability and concurrent validity, we especially recommend study of the extent to which these methods are useful for iteratively informing intervention and implementation modifications during a project. Future studies could also evaluate the value and cost-effectiveness of these brief, pragmatic assessment methods compared with more traditional evaluation procedures. Future research is indicated that helps inform the overarching question of which assessment methods are most useful in what settings.

ETHICS STATEMENT

This study was not considered research per VHA ORO policy 1058.05.

AUTHOR CONTRIBUTIONS

BR conceptualized and prepared the first draft of the paper, developed the assessment methods and tool, led the interpretation of the data interpretation, and reviewed the final draft of the paper. MM participated in the conceptualization and development of the first draft of the paper, participated in the
development of the assessment methods and tools, led the data collection and data analysis, participated in the interpretation of the data, and reviewed the final draft of the paper. CB participated in the conceptualization and drafted sections of the paper, participated in the development of the assessment methods and tools and the data interpretation and reviewed the final draft of the paper. RA participated in the development of the first draft of the paper, the development of the assessment methods and tools, data collection, data analysis, and interpretation of findings, and reviewed the final draft of the paper. RB participated in the drafting of sections of the paper, serves as the PI for the Multimodal Pain project, participated in the interpretation of the data and reviewed the final draft of the paper. RG co-led the conceptualization of the paper with BR, developed the first draft of the paper with BR, developed the original assessment tool, participated in the development of the revised assessment methods and tool, participated in the data interpretation, and reviewed the final draft of the paper.

**FUNDING**

Funding for this paper was provided through the VA Quality Enhancement Research Initiative (QUERI) Program (Triple Aim QUERI Program, QUERI 15-268) and the Office of Rural Health in partnership with the Office of Nursing Services through the Enterprise-Wide Initiative (EWI) (Rural Transition Project).

**REFERENCES**


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer SY and the handling Editor declared their shared affiliation.

Copyright © 2018 Rabin, McCreight, Battaglia, Ayele, Burke, Hess, Frank and Glasgow. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
The Family Check-Up 4 Health (FCU4Health): Applying Implementation Science Frameworks to the Process of Adapting an Evidence-Based Parenting Program for Prevention of Pediatric Obesity and Excess Weight Gain in Primary Care

Justin D. Smith1,2,3, Cady Berkel4,5, Jenna Rudo-Stern4, Zorash Montaño6, Sara M. St. George7, Guillermo Prado7, Anne M. Mauricio4, Amanda Chiapa8, Meg M. Bruening9 and Thomas J. Dishion4,10

1 Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, United States, 2 Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States, 3 Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, IL, United States, 4 REACH Institute, Department of Psychology, Arizona State University, Tempe, AZ, United States, 5 Phoenix Children’s Hospital, Phoenix, AZ, United States, 6 Children’s Hospital of Los Angeles, University of Southern California, Los Angeles, CA, United States, 7 Department of Public Health Sciences, University of Miami Miller School of Medicine, Miami, FL, United States, 8 Yale Child Study Center, New Haven, CT, United States, 9 Department of Nutrition, Arizona State University, Tempe, AZ, United States, 10 Oregon Research Institute, Eugene, OR, United States

Implementation experts have recently argued for a process of “scaling out” evidence-based interventions, programs, and practices (EBPs) to improve reach to new populations and new service delivery systems. A process of planned adaptation is typically required to integrate EBPs into new service delivery systems and address the needs of targeted populations while simultaneously maintaining fidelity to core components. This process-oriented paper describes the application of an implementation science framework and coding system to the adaptation of the Family Check-Up (FCU), for a new clinical target and service delivery system—prevention of obesity and excess weight gain in primary care. The original FCU has demonstrated both short- and long-term effects on obesity with underserved families across a wide age range. The advantage of adapting such a program is the existing empirical evidence that the intervention improves the primary mediator of effects on the new target outcome. We offer a guide for determining the levels of evidence to undertake the adaptation of an existing EBP for a new clinical target. In this paper, adaptation included shifting the frame of the intervention from one of risk reduction to health promotion; adding health-specific assessments in the areas of nutrition, physical activity, sleep, and media parenting behaviors; family interaction tasks related to goals for health and health behaviors; and coordinating with community resources for physical health. We discuss
the multi-year process of adaptation that began by engaging the FCU developer, community stakeholders, and families, which was then followed by a pilot feasibility study, and continues in an ongoing randomized effectiveness-implementation hybrid trial. The adapted program is called the Family Check-Up 4 Health (FCU4Health). We apply a comprehensive coding system for the adaptation of EBPs to our process and also provide a side-by-side comparison of behavior change techniques for obesity prevention and management used in the original FCU and in the FCU4Health. These provide a rigorous means of classification as well as a common language that can be used when adapting other EBPs for context, content, population, or clinical target. Limitations of such an approach to adaptation and future directions of this work are discussed.

Keywords: adaptation, implementation strategies, family check-up, family check-up 4 health, obesity prevention, primary care, scaling out

INTRODUCTION

Translation of evidence-based interventions, programs, and practices (EBPs) for children and adolescents to the real-world service systems that can support them is a challenging endeavor and the lack of wide scale dissemination and implementation is well documented (1, 2). EBPs grounded in the principles of parent training are highly effective at preventing a host of common mental and behavioral problems in youth (3) and have been found to be effective when tested under more “real-world” conditions (4). That is, conditions more closely aligned to typical operations and resources available in non-research settings. Parenting programs are slowly making their way into the service delivery systems where youth and families are served. These include social services, schools, and home visitation. A relevant setting where such interventions have not largely been adopted is pediatric primary care. This setting is particularly relevant for preventive parenting interventions as the majority of children in the U.S. receive annual primary care services (5); low-income children have high rates of access (6); parents expect to receive parenting advice from physicians and view them as respected experts; there are potentially stable mechanisms to fund these EBPs, whereas in other settings, these are lacking; and this setting does not hold the stigma that others, such as schools, do (7). Parenting in general, and the effects of parenting interventions specifically, are linked to both mental and physical health conditions, making these programs highly relevant as a primary prevention strategy for improving the health of all children (8). The existing primary care system is the ideal context for parenting interventions to be implemented. One of the barriers to doing so, however, is the need to adapt parenting programs for the primary care context and the populations that would receive these interventions.

Use of adaptation as an implementation strategy is common and aimed at making EBPs more appropriate, feasible, cost-effective, and acceptable for both the target population and service delivery system (9). Articles describing adaptations for particular populations are more common in the literature than are those for delivery through a new delivery system. Most common population adaptations are for different cultural groups. Appreciation for these adaptations grew as problems with focusing behavioral interventions exclusively on the majority culture, typically non-Latino White families, emerged. Specifically, focusing only on the majority culture often led to the EBP being ineffective with, or simply unpalatable to, culturally diverse populations (10–12). Adaptation to a new delivery system involves pursuing an alternative means through which to reach the target population. This form of adaptation has traditionally been done by changing the context through which an intervention is delivered (e.g., from schools to mental health or social service systems).

A traditional assumption in the field is that when EBPs are adapted, they need to be rigorously re-tested to ensure positive effects of the original program are not degraded. However, Chambers et al. (13) and others have argued that adaptation of EBPs—done in a way that maintains fidelity to the core components—should result in at least comparable effect sizes and are, perhaps more importantly, likely to be sustained. The results of a recent review of adapted EBPs by Wiltsey-Stirman et al. (14) found little evidence that adaptations were detrimental to effectiveness. Relatedly, however, they also found limited consistent evidence that adapted protocols outperformed the originals; the exception was the addition of components, which had a modest positive impact on outcomes.

A process of "scaling out" has been recommended to more rapidly increase the reach of EBPs (15). Scaling out differs from the more common practice of scaling up an EBP, which means to spread to additional units of the same or a very similar context, and customarily targeting the same population, for which the EBP was originally tested and shown to be effective. When scale up occurs, there is an assumption that the EBP will be delivered in the same way to the same type of population or people and, therefore, health benefits will align with previous research if there

Abbreviations: ASU, Arizona State University; CAB, community advisory board; CDC, Centers for Disease Control and Prevention; CHIP, Children's Health Insurance Program; CORD, Childhood Obesity Research Demonstration project; EBPs, Evidence-based programs and practices; FCU, Family Check-Up; FIT, Family Interaction Task; FCU4Health, Family Check-Up 4 Health; PCH, Phoenix Children's Hospital.
is sufficient fidelity (16). It is thought that adaptation of the EBP either is unnecessary or will simply not occur in a meaningful way in a scale up scenario. In contrast, scaling out is defined by Aarons et al. (15) as a deliberate effort to adapt an EBP and broaden its delivery (a) to a different delivery system, but with the same target population as previous trials; (b) to a different target population, but within the same delivery system as previous trials; or (c) to a different target population and through a different delivery system than those of previous trials.

There are a number of process frameworks to prospectively guide the adaptation of EBPs, some of which are more generic (17, 18), while others are specific to cultural adaptation [see Barrera et al. (10)] or to technology-based platforms (19).

Adapting an EBP for a new clinical target outcome goes beyond these models in important ways and is least represented in the literature. Aarons et al. (15) suggest that the “new target population” refers to the characteristics of the population, such as developmental period (i.e., age), culture, or socioeconomic status, but that the clinical target outcome is typically the same (e.g., a preventive intervention targeting problem behaviors in young children vs. adolescents). There are instances, however, when adapting an EBP for a new clinical target outcome may be warranted. The evidence for doing so may come from a number of potential sources, including studies examining collateral benefits of an intervention (i.e., effects on outcomes not directly targeted). For example, the Familias Unidas program was originally designed to prevent and reduce behavior problems and substance use in Latino adolescents, but it has also had positive effects on adolescents’ internalizing symptoms and suicidal behaviors (20, 21). Another common collateral effect of parenting programs is improvement in parental mental health, such as reducing parents’ depressive symptoms [e.g., Beach et al. (22), Shaw et al. (23)].

This article uses concepts and frameworks from the field of implementation research to present and document the process of scaling out an evidence-based parent training program for a new clinical target and service delivery system. The Family Check-Up [FCU; Dishion et al. (24)], which was originally tested in public schools, community mental health clinics, and home visiting services for families with youth at risk for problem behaviors, has been adapted to target the prevention of obesity and excess weight gain in collaboration with the primary healthcare system. This paper attempts to accomplish three aims: First, we propose four levels of evidence (minimum, preferred, preferred plus, and optimal) as a framework to guide decision-making around the adaptation of an EBP for a new clinical target—and this is not represented in the adaptation literature. These levels pertain to the justification for conducting this type of adaptation. Second, we categorize the modifications and adaptations made to the FCU based on an existing framework, which was selected because it was developed in the context of implementation science, is comprehensive, and can be applied retrospectively (25). We describe our process by detailing the various methods and activities that were used to obtain salient guidance. These included analyses of existing data and reviews of the literature by the academic team, research-practice partnerships with local agencies, and collaboration with diverse community stakeholders. Activities with stakeholders comprised formal and informal meetings, a pilot study at a partner agency, establishing and regularly convening a community advisory board (CAB), and conducting a multisite randomized trial (currently ongoing) called the Raising Healthy Children study1 (26). Last, we apply a recent standardized taxonomy for specifying the behavior change techniques used in behavioral interventions for pediatric obesity (27) to the resulting adapted and enhanced version of the FCU, which we call the Family Check-Up 4 Health (FCU4Health), and contrast that with the original program components. This step is important in demonstrating that FCU4Health aligns with the characteristics of other EBPs for the prevention of pediatric obesity and excess weight gain. We anticipated that the process of adapting FCU for obesity prevention in primary care would center around changes and modifications to the content of the intervention to more specifically target weight-related variables and modifications to the delivery of the program to better align with the context of primary care, specifically aligning with the national recommendations for the prevention of excess weight gain and a staffing model consistent with the primary healthcare system.

Proposed Levels of Evidence for Adapting an EBI for a New Clinical Target

Three of the authors (Smith, St. George, Prado) developed the proposed four levels of evidence to consider when endeavoring to adapt an EBP for a new clinical target (see Figure 1). The need to develop these levels of evidence emerged as these authors considered making adaptations for a new clinical target, which differs from adaptations for a new population or setting. With the large body of evidence indicating collateral effects of parenting interventions [see Van Ryzin et al. (28)], such a guide for adapters of these programs specifically for new clinical targets would be useful. Each level is cumulative; it requires the newly specified set of evidence in addition to the evidence listed in each of the previous levels. Although not necessary, it would be preferable each level of evidence be documented within the target population (e.g., Latino immigrants). If research with a specific target population is not available, this should not necessarily limit the adaptation of the EBP for the new clinical target. In situations where evidence in the target population is unavailable, researchers may want to consider whether (a) theory or input from relevant stakeholders and (b) cross-sectional OR (preferably) longitudinal research support the causal relations between the program, mechanisms of action, and the new clinical target. Experimental designs, such as randomized trials, are preferred at each level. Other designs (e.g., pre-post) are acceptable but multiple studies with consistent significant relations would be needed. In our descriptions of each level, we integrate information from our work with FCU as illustration. However, evidence can be garnered from different EBPs that have similar intervention strategies and theories of action (see Level 3 for an example of drawing from other EBPs to support adaptation of FCU).

1 The FCU4Health as described in this article is being tested in the Raising Healthy Children study compared to primary care services as usual.
Level 1: Minimum Evidence

To consider adapting an EBP for a new clinical target, it is important first to determine whether there is sufficient evidence demonstrating that the mechanisms of action (e.g., family functioning) of the EBP are related to the new clinical target (e.g., physical activity, weight loss). Such evidence would require that the EBP’s mechanism(s) of action have been shown to influence the clinical target in more than one cross-sectional study or at least one longitudinal study. For example, family functioning has been found to be related to childhood obesity and obesity-related outcomes in both cross-sectional (29) and longitudinal research (29–31).

Level 2: Preferred Evidence

This level requires all the evidence listed in Level 1 and documented evidence that the EBP impacts the mechanism(s) of action. For example, if a parenting intervention targeting substance use is being considered for adaptation to obesity or obesity-related outcomes, that EBP should have documented evidence that it leads to improvement in the mechanism(s) of action. Mediational analyses of randomized trials have demonstrated that the FCU prevents and reduces youth substance use through improvements on the same family processes that have been linked to obesity and obesity-related outcomes in longitudinal studies [e.g., (24, 32–38)].

Level 3: Preferred Plus Evidence

This level requires the criteria listed above and evidence that the EBP has an impact on the new clinical target. There are a few examples of effects of parenting programs on obesity. For example, Brotman et al. (39) found that a parenting intervention not focused on improving physical health significantly reduced body mass index (BMI) 5 years post-intervention. (40) provide a review of similar effects of parenting programs on obesity.

Level 4: Optimal Evidence

This level requires the previous criteria plus evidence that the mechanism of action mediates the effects of intervention on the new clinical target. The original version of FCU, which was not designed to target obesity and obesity-related outcomes, has had collateral effects on obesity in two randomized clinical trials. In early childhood, effects on weight gain trajectories were mediated by immediate improvements in observed positive behavior support skills, which were in turn related to serving...
Adapting the Family Check-Up for the Prevention of Obesity and Excess Weight Gain

The components and content of the original FCU model are described in Table 1. Additional information is available in Dishion and Stormshak (45) and Smith (46). In brief, the FCU involves a 3-step process comprising an initial interview with the family, an ecological family assessment (multimethod, multirater), and a motivation-enhancing feedback session. During the feedback session, family strengths, and areas for potential intervention identified in the ecological assessment are discussed with the caregiver(s) and motivational interviewing is used to motivate families to make change and engage in additional intervention. The primary form of subsequent intervention is behavioral parent training and a variety of community-based support services for the child (e.g., individual mental health intervention) and the caregiver(s) (e.g., marital or substance abuse counseling).

Although adaptation of the FCU for the prevention of childhood obesity and excess weight gain and the primary care context has not been described previously, the program has been previously adapted for various populations and delivery systems. Figure 2 provides a schematic of the various adaptations of the FCU and the approximate chronology of these efforts. A critical facet of each adaptation is retention of the core components of the program and intervention strategies targeting age-appropriate parenting and behavior management skills [for further discussion, see Smith and Dishion (47)]. FCU was originally designed for the prevention of problem behaviors in the transition from middle to high school that increase the risk of substance use, high-risk sexual behaviors, violence, and other related outcomes (48) based on the successful Adolescent Transitions Program (49) and the Parent Management Training Oregon Model (50). The initial trials of the FCU occurred in public middle schools (children age 12 to 14 years) in underserved urban areas [see (35), (51)]. As described by Smith et al. (38), the FCU was designed to be multiculturally responsive and empirical studies have shown that different racial and ethnic groups participate in and benefit from the program similarly. However, there was not a specific adapted version of FCU for each racial/ethnic group; rather, the program was individually-tailored to the specific needs of each family [see Smith et al. (38)]. The first adaptation of FCU for a specific racial/ethnic group was for American Indian youth (52). Next, the program was adapted for families with young children ages 2 and 12 years (24) and for delivery through home visitation rather than embedded in schools (53). Then, FCU was adapted for delivery within community mental health clinics (54). More recently, the original FCU was adapted for delivery within and in coordination with primary healthcare systems (55). Ongoing studies are testing the effectiveness of (a) a version of the program adapted for emerging adults (ages 19 to 23 years) and their parents (56) and (b) an Internet-based delivery of FCU to families in rural areas identified in middle schools (57). In each of these situations, the context of FCU delivery and the population were targeted for adaptation, but the primary clinical target (i.e., reduction of child problem behaviors through the improvement of family management) remained consistent.

The developer of the FCU, Thomas Dishion, and other FCU researchers at multiple institutions have recently undertaken efforts to adapt the program to fit better within and in coordination with pediatric primary care. These efforts coincide with a national movement to implement evidence-based parenting programs within primary care for the prevention and the treatment of behavioral and mental health conditions (7, 8, 58). In addition to the effort described in this paper, Shaw et al. (59) have been working in partnership with primary care practices to reach children in need of family support services to prevent substance use in pre- and early adolescence and improve school readiness in young children. Their approach, however, involves less adaptation to the FCU itself as they identify eligible families in primary care, but then deliver the FCU through the previously successful home visitation model (outside of the primary care office). This linking of an EBP with the primary care context is important, but places fewer demands on the system to adopt the FCU compared to a more integrated delivery model, and therefore, there is lesser need to adapt the FCU for the demands of the context and is potentially more quickly translated. Evaluating this hypothesis is a primary aim of the ongoing Raising Healthy Children study (26) and a second study also being led by Smith and Berk that is funded by the United Department of Agriculture (Grant number: 016–10799). Relatedly, Polaha and colleagues pilot tested the FCU in primary care clinics for young children’s mental and behavioral concerns. The process and outcomes of these efforts are discussed in Smith et al. (60) and Smith and Polaha (55).

In addition to adapting the FCU for delivery in the primary care context, we also undertook a process of enhancing the program to better target health behaviors and parental supports to prevent obesity and excess weight gain. The findings of previous research, including our own with the original FCU, have resulted in efforts to enhance evidence-based parenting programs to more specifically address behaviors related to maintaining a healthy weight. Familias Unidas (St. George, S. M., Messiah, S. E., Sardinas, K. M., Poma, S., Lebron, C., Tapia, M.,....
TABLE 1 | Classifications of CONTEXT modifications to the original FCU in developing the FCU4Health program.

<table>
<thead>
<tr>
<th>What is modified?</th>
<th>FCU4Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format</td>
<td>1-on-1</td>
</tr>
<tr>
<td>Health maintenance model: Annual feedback sessions with individually-tailored family support and referral</td>
<td>1-on-1: Intensive: 3 feedbacks in 6 months with individually-tailored family support and referral (Goals: USPSTF recommendation of 25–50 h)</td>
</tr>
<tr>
<td>Setting</td>
<td>Schools</td>
</tr>
<tr>
<td>Community mental health</td>
<td>Pediatric primary care (i.e., behavioral health service, integrated care)</td>
</tr>
<tr>
<td>Social services</td>
<td>Home visitation</td>
</tr>
<tr>
<td>Personnel</td>
<td>Master’s and doctoral-level mental health providers (e.g., social workers, psychologists)</td>
</tr>
<tr>
<td>Referral: school, mental health provider, parent</td>
<td>Referral: pediatrician</td>
</tr>
<tr>
<td>Population</td>
<td>Families with youth at risk for problem behaviors (e.g., oppositional, substance use, high risk sex, school failure)</td>
</tr>
<tr>
<td>Ages 2–17 years (across multiple trials)</td>
<td>Families with youth at risk for obesity and excess weight gain (e.g., poor diet, low physical activity, racial/ethnic/cultural groups with disproportionate risk)</td>
</tr>
<tr>
<td></td>
<td>Low-income, underserved</td>
</tr>
<tr>
<td></td>
<td>Ages 6–12 in Raising Healthy Children</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
</tr>
</tbody>
</table>

Prado, G. *Familias Unidas por la salud & bienestar: Adapting an evidence-based substance use and sexual risk behavior intervention for obesity prevention in Hispanic adolescents*, submitted for publication) and Lifestyle Triple P (61) are other examples in the family intervention literature of adaptation for this new clinical target. The activities and sources of data used in the adaptation and enhancement of the FCU for the prevention of excess weight gain through primary care are presented in the Method section and how each contributed to the adapted and enhanced version of the program—the FCU4Health—is presented in the Results section.

METHODS

Procedure
Adapting an EBP when scaling out should generally comprise an iterative, multi-method, and multi-informant process. Our process occurred through a variety of activities and sources of data. These are described in the chronological order in which they occurred. Continuation or repetition over time is noted as appropriate. This study was carried out in accordance with the United States Department of Health and Human Services (HHS) policy for the protection of human subjects. The protocols of the pilot study and ongoing *Raising Healthy Children* study from which data were drawn for this article were approved by the institutional review boards of Arizona State University and the Phoenix Children’s Hospital. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

Evidence From Prior Trials of FCU For Adolescents
As summarized above and in Figure 1, there was optimal evidence for adapting the FCU to focus on health behaviors. Enhancement for the prevention of obesity and excess weight gain was informed by analyses of prior trials of the FCU showing mediated effects of the intervention on obesity and related processes (41, 43).

Meetings With Pediatricians And Social Workers at Phoenix Children’s Hospital (PCH)
In 2011, FCU developer Dishion and collaborator Berkel began a series of formal and informal meetings with pediatricians and social work staff at PCH concerning the acceptability and appropriateness of using the FCU in general pediatrics. These meetings comprised presentations by Dishion and sharing of FCU materials. The chief of pediatrics and other clinic leadership also attended. This partnership was made possible by Berkel’s dual appointment at ASU and PCH’s general pediatrics care coordination program (Berkel, C., Araica, E., Smith, J. D., Tovar-Huffman, A., Beaumont, S. W., & Shaw, T. *Connecting families: Implementation and outcomes of a comprehensive care coordination program*, manuscript in preparation).

Pediatrician Needs And Attitudes Survey
As a result of these meetings focused on exploring use of FCU in general pediatrics, in 2012, Berkel et al. (62) conducted a survey of the 20 physicians in the general pediatrics clinics about their concerns for families and attitudes toward implementing the FCU in the clinic. The top three areas of concern were obesity (100%), nutrition (95%), and parenting (90%). All respondents perceived a need for a program like the FCU that could also address family factors related to weight management. Open-ended responses, provided by 70% of respondents, reflected themes of limited time to convey important, tailored health information; the desire to increase parent understanding and empowerment to support children’s health behavior change; and a recognition of the many barriers to families’ being able to follow through with recommendations for healthy lifestyle behavior.
change. Further, pediatricians reported feeling unprepared to contend with these family-level barriers to follow through due to their lack of training in this area and their many practice demands. In 2011, Dishion and Berkel submitted a grant to test the effectiveness of the FCU when implemented by general pediatrics care coordinators in early childhood (birth to five). This grant was not funded. In 2014, Berkel, Dishion, and Smith submitted a grant to test the effectiveness of the FCU through care coordination with adolescents, and Smith, Marisol Perez, and Dishion submitted a grant to test FCU as an add-on parenting support service for families in an outpatient specialty care clinic for obesity in the hospital. Neither of these grants were funded.

**Pilot Trial of FCU in Pediatric Primary and Specialty Care**

Dishion was awarded a seed grant from Arizona State University (ASU) to conduct a pilot feasibility trial of implementing the FCU in general pediatrics and a specialty clinic in the department of gastroenterology for children with non-alcoholic fatty liver disease at PCH. These two clinics were selected to obtain information about feasibility, appropriateness, and acceptability of FCU for outpatient primary and specialty care, and how the different population characteristics could inform adaptation of the program to better address prevention of obesity and excess weight gain across levels of disease progression. The pilot trial was run by Dishion, with assistance from Berkel, Smith, and ASU clinical psychology graduate students (Montaño, Rudo-Stern, Chiapa) who provided the FCU. Eleven families and fourteen families were consented and participated in some aspect of FCU in the fatty liver and general pediatrics clinics, respectively. The project demonstrated a need to adapt the FCU to fit clinic procedures, which necessitated adaptations to program delivery and addition of a relevant screening process and instrument (63). Activities followed a participatory research approach with health care staff (pediatricians, nurses, dieticians), patients, and their families. All eligible families were offered the FCU at no charge and received a $20 gift card for completing the assessments. Qualitative findings from stakeholder interviews with pediatricians ($n=11$) and dieticians ($n=4$) indicated that these stakeholders: (1) desired to involve families in the FCU; (2) actively offered FCU to families; (3) saw a need for...
an intervention to help families with parenting practices as they relate to children’s physical health; and (4) felt the name “Family Check-Up” is appropriate in this setting. Additionally, the average score on the Evidence Based Practice Attitudes Scale (64) reported by physicians and dietitians (n = 15) was a 44 (out of 50) indicating adequate acceptability. Results of family interviews revealed (1) receptivity to the family-centered nature of the FCU to support parents; (2) parents see themselves as the primary source of support for their children; (3) acceptability of a program that supports parents in implementing treatment recommendations; (4) if an intervention that provided this type of additional support for parents was being offered in their clinic, they would find it appealing and would enroll; and (5) the name “Family Check-Up” is appropriate in this setting. Individual outcomes were not assessed as part of this trial as the goal was determining feasibility and needed adaptations and enhancements.

Raising Healthy Children Study: Hybrid Effectiveness-Implementation Trial in Primary Care

In 2014, Smith, Perez, Dishion, and others submitted a proposal for federal funding to conduct a randomized trial of the FCU in specialty care for the management of obesity. The proposal was not funded. In 2015, Berkel and Smith led a proposal in response to a request for applications issued by the Centers for Disease Control and Prevention under the Childhood Obesity Research Demonstration (CORD) projects, version 2.0 announcement. The proposal was selected for funding and the Raising Healthy Children study officially began on June 1, 2016. Specific Aim 1 of the project was stated as “Finalize the adaptation of the Family Check-Up 4 Health (FCU4Health) program, which was initially adapted and piloted in pediatric primary healthcare, based on input from a CAB and partner clinics.” In brief, the trial is an effectiveness-implementation hybrid trial comparing an integrated/co-located model of care with a referral to external services model in partnership with three primary care agencies. Families are randomly assigned to FCU4Health (n = 200) or to clinic services as usual plus information (n = 150). The FCU4Health protocol in the Raising Health Children study consists of three family health behavior assessments and three feedback sessions, with individually-tailored family support sessions and referral to community-based services, over a 6-month period. An assessment 1 year after baseline will be used to examine lasting effects. The full study protocol is available in Smith et al. (26). As of this writing, the trial is still enrolling participants and providing the FCU4Health to families randomized to that arm. FCU4Health coordinators and supervisors meet regularly to discuss barriers and refinements to delivery process.

Community Advisory Board

The CORD 2.0 funding mechanism entailed the inclusion of a CAB with the goal of ensuring that at the conclusion of the project, the program would be ready for dissemination. The request for applications stated, “Collaboration with state CHIP and/or Medicaid offices to advise a state-wide or regional level project and to be part of a stakeholder group that can help generate suitable recommendations for sustainability and program components to be further replicated or scaled.” In addition to a representative from the state Medicaid office, known as the Arizona Health Care Cost Containment System (AHCCCS), our CAB includes leadership and direct service providers from local agencies (including our partner agencies for the project); stakeholders from relevant local entities (e.g., local health department, Arizona chapter of the American Academy of Pediatrics); representatives from health insurance plans; and researchers in obesity prevention, nutrition, and health disparities. We used community engaged dissemination and implementation research methods (65) to inform the conduct of our research and the execution of implementation.

We convened our first official CAB meeting in May 2016 (3h) to prepare for the project starting. We next convened a day-long CAB meeting in September 2016, in which we held three concurrent work groups with the aim of obtaining guidance on three key aspects of the project: (1) evidence needed for post-project adoption; (2) integration of FCU4Health into the pediatric primary care system; and (3) program components to increase effectiveness for prevention of excess weight gain. Clearly, the latter two are directly relevant to adapting the program’s delivery and enhancing its content. Berkel et al. (Berkel, C., Rudo-Stern, J., Villamar, J., Wilson, C., Flanagan, E., Smith, J.D. Recommendations from community partners to promote sustainable implementation of evidence-based programs in primary care, manuscript in preparation) discuss our partnership formation and the products of the CAB through the qualitative analysis of these work groups. Relevant findings for adaptation and enhancement are presented in the Results section. Recently, the CAB and four collaborators from the Obesity Prevention and Control Branch at the CDC, assembled for a 3 h meeting in September 2017 for updates on project progress and a discussion of successes and challenges to achieving the stated aims.

RESULTS

This section of the paper describes the FCU4Health program and classifies in what ways the original FCU was adapted for primary care or enhanced for the prevention of obesity and excess weight gain. Importantly, we also note important aspects of the FCU that were not changed for FCU4Health; these were critical for maintaining fidelity to the core components of the program and the underlying theory of change. We use the framework and coding system for modifications to EBPs developed by Stirman et al. (25). There is also a description of the activities (listed in the Method section) that were used in making the described adaptations. Finally, we use the JaKa et al. (27) taxonomy to specify the behavior change techniques used in FCU4Health as a means of providing a standardized comparison to similar programs.

Adaptation and Enhancement

The Stirman et al. (25) framework considers first the type of modification: content of the EBP or context in which it is being delivered. Within content modifications, 12 categories
concerning the nature of the modification are identified and the level at which the modification occurs is specified (e.g., individual patient, clinic population). For context modifications, 5 categories were identified and include changes to the format, the setting, or the patient population (that do not result in changes to the actual content of the EBP). For each type, who was responsible for the modification is also indicated. Table 1 presents our classifications of the context modifications and Table 2 the context modifications made to the original FCU in development of the current model of FCU4Health. The narrative that follows in this section is intended to supplement and synthesize the information in the table by providing information on the timing of the modification and the sources of data that were used. This is one aspect of the Stirman et al. framework where we diverge. As it was intended to be applied retrospectively and not prospectively, we note when modifications were a priori (by the program developers) or after data sources indicated a need. The narrative also covers a third area of modification in the Stirman et al. framework: training and evaluation.

**Context Modifications: Adaptation for Primary Care**

The 1-on-1 delivery format of the FCU and use of home visitation during early childhood was retained for FCU4Health. The motivation for and merits of a 1-on-1 approach that occurs largely in the family home, compared to the common group-based delivery of parenting programs at a central location, are discussed in Smith et al. (53) and Dishion and Kavanagha (48). Home visiting for delivery of behavioral health is particularly germane to coordinating with pediatric primary care due to the space limitations of typical medical offices for use by behavioral health staff. In FCU4Health, identification, referral, and initial contact (ideally) occur in the primary care office and the remaining intervention services predominantly occur in the family home or a community location (e.g., community center, YMCA). However, the delivery strategy is flexible and is currently being done in multiple ways in an ongoing trial aligning with the staffing, space, and preference of the clinics involved (Berkel et al., submitted). One major format modification that was made for the Raising Healthy Children study concerns the intensity of services provided. The original FCU was designed for selected and indicated prevention and was intended to be delivered using a health maintenance approach. Specifically, each year the family has a comprehensive assessment and a “feedback session” to build motivation and plan follow-up services for which the intervention intensity (number and frequency of sessions) is guided by the current level of need (66). In this project, a more intensive model of delivery is being used. The CORD 2.0 RFA required a delivery approach that would meet the recommended

| TABLE 2 | Classifications of CONTENT modifications to the original FCU in developing the FCU4Health program. |
|-------------------------------|-----------------------------------------------------------------------------------|-----------------|-----------------|
| FCU                                                                                     | FCU4Health                                                                         | By whom were modifications made? | Level of delivery | Nature of the modification |
| Screening—child behaviors and family risk factors for ineffective parenting              | Screening—child body mass index (BMI)                                              | Program developersResearchers (required by funding agency) | System level      | Substituting               |
| 3 contacts (initial interview; ecological family assessment; feedback session)         | 2 contacts (combined initial interview and family health routines assessment; feedback session) | Program developersResearchers (required by funding agency) | System level      | Shortening/condensing      |
| Ecological assessment surveys—focused on ecological influences on children’s adaptations and behavior and on parent’s ability to manage the family | Family health routines surveys—added health routines and behaviors module (e.g., dietary practices, meettime and sleep routines, physical/ sedentary activity, health related quality of life) | Program developersResearchers       | System level      | Adding elements            |
| Family Interaction Tasks—focus on risk reduction of factors related to problem behaviors (e.g., monitoring, limit setting) | Family Interaction Tasks—focus on promoting healthy goals/behaviors and setting limits on unhealthy behaviors | Program developersResearchers (required by funding agency) | System level      | Substituting Tailoring/tweaking/ refining |
| 5 tasks, 5 min each                                                                  | 3 tasks, 3-4 min each                                                              | Coalition of stakeholders         | System level      | Adding elements            |
| N/A Referrals to community services and supports                                       | Anthropometric evaluation (BMI, body composition) of child and other family members, who are encouraged to provide this data | Program developersResearchers | System level      | Tailoring/tweaking/refining |
| No explicit focus on nutrition or health behaviors related to obesity and excess weight gain | Programs and services for diet, nutrition, physical activity and services to address social determinants of health | Program developersResearchers     | System level      | Adding elements            |

Frontiers in Public Health | www.frontiersin.org 144 October 2018 | Volume 6 | Article 293
25 to 50 h of intervention time over a 6-month period specified by the US Preventive Services Task Force for youth with a BMI for age and gender of ≥85th percentile (67). This requirement led us to devise a condensed health maintenance approach with three feedbacks in 6 months (Months 1, 3, 6), rather than the customary annual feedback, to facilitate achieving the hourly target, allow us to continually tailor the intervention for each family, and explicitly address motivation to change behavior—a primary challenge in family-based intervention for the prevention of excess weight gain (68) and an explicit target of the FCU and FCU4Health. This schedule aligns with the suggested frequency of visits to primary care for children with obesity (69). In this way, the FCU4Health is being delivered as an indicated intervention in the Raising Healthy Children study. However, in an ongoing trial of FCU4Health funded by the United States Department of Agriculture to Berkel and Smith (Grant number 016–10799), it is being delivered as a selected intervention for young children (ages 2 to 8 years) who screen positive for poor dietary habits but who do not have an elevated BMI. Rather than the intensive delivery of the program as is being done in the Raising Healthy Children study, delivery of FCU4Health occurs annually for 3 consecutive years with individually-tailored intervention plans (i.e., number of hours each year vary from 3 to 10) to correspond with each child and family’s specific level of need.

Concerning the setting, we have previously discussed our scale-out effort to primary care. The program developers sought to take the FCU into this service context for a number of reasons. First, it is a setting that serves a high proportion of children and families; parents are typically present at children’s healthcare visits; and parents are used to receiving advice from pediatricians as a trusted source of information (7). These factors generally support a parenting intervention in primary care. Specific to shifting our clinical target of obesity and excess weight gain, primary care is a context where weight and weight-related behaviors are thoroughly embedded, it is the only system that regularly tracks weight throughout childhood, and parents may be more receptive to learning about their child’s risk for obesity from their pediatrician than in other contexts (7) where identifying children with elevated BMI creates concerns about confidentiality and stigma (70). Our early and ongoing meetings with stakeholders, survey of pediatricians’ needs, and pilot trial provided the necessary evidence that such a program is acceptable and appropriate for this setting. Formal data collection on acceptability and appropriateness is ongoing in the Raising Healthy Children study, but no major concerns have emerged up to this point.

The personnel that typically deliver FCU were largely maintained for FCU4Health. The primary providers are Master’s-level clinicians with backgrounds in mental and behavioral health. In working with our partner agencies, and discussing children’s primary healthcare practices more broadly with the CAB, we elected to allow professionals from obesity-related fields, such as health promotion, nutrition, and public health, to be trained to deliver the intervention, as these are the professional roles that serve similar functions to FCU4Health in pediatric healthcare agencies and often have training in motivational skills. Further, some components of the FCU4Health (i.e., conducting the assessment, connecting families with referrals to community resources to address contextual needs) may be completed by community health workers or promotoras. This diffusion of responsibilities fits with the medical home framework in which each person in the clinic performs roles in accordance with their training and abilities (71). The procedures for implementation vary, however, by the agency or clinic depending on their available personnel and other resources and the model of behavioral health services used (e.g., integrated care, coordinated care, colocation, referral to external service provider). Thus, when implementing FCU4Health, there is a need to accomplish specific program activities but the manner in which this is done, who is responsible, and even where they are delivered—in the clinic, the home, or another agency’s offices—is flexible (72; Berkel et al., submitted).

With the change of setting came a change in the referring professional. In previous trials, referrals originated with school personnel, the parents, or a mental health provider. In keeping with typical procedures in pediatric primary care, which was also a requirement of the CORD RFA, the pediatrician identifies children with elevated BMI and refers to the FCU4Health. In our pilot trial and in meetings with our partner clinics and the CAB, this procedure was found to be feasible and appropriate.

Population modifications centered on the new clinical target: pediatric obesity. Instead of the FCU procedure of targeting characteristics of children and families focused on risk reduction for problem behaviors, FCU4Health targets families with youth at risk for obesity and excess weight gain, but with a health promotion approach. Characteristics considered include behavioral risk factors, such as poor dietary practices and low physical activity, and also membership in sociodemographic groups that are disproportionately affected by the obesity epidemic, including low-income and racial/ethnic minority families (73). The age of the targeted youth is intended to be the same as the original FCU, which is 2 to 17 years, however, CORD 2.0 funding is for inclusion of children ages 6 to 12 years.

In summary, nearly all context modifications were made by the program developers, who are also the researchers on the project. Some modifications were directly influenced by the CORD 2.0 RFA. Agency administrators (i.e., leadership at our partner clinics) were influential in the decision to include related professionals in the delivery of FCU4Health. Nearly all of these modifications were determined a priori to the grant proposal, based on our prior experiences (e.g., meetings, pilot trial).

**Content Modifications: Enhancement for Pediatric Obesity**

Content modifications to FCU primarily involved tailoring and adding elements to address obesity and health behaviors. Care was taken to retain core components of the FCU in order to maintain its effectiveness at improving parenting and family functioning, which we found mediated effects of the program on obesity in childhood and adolescence to adulthood (41, 43). Content modifications were made in close collaboration with our CAB. Berkel et al. (manuscript in preparation) report the primary qualitative results of an analysis of transcripts from three working groups conducted at our September 2016 CAB meeting. These working groups discussed the topics of (1) fit...
of the FCU4Health within primary care, (2) components of the program for prevention of obesity and excess weight gain and management of co-occurring concerns, and (3) evidence needed to support sustainment of the program after the trial. Salient results from this qualitative research are included in the following sections.

The procedures for delivering FCU4Health differs somewhat from FCU due to the demands of the delivery setting and the new clinical target. Members of our CAB engaged in a working group on the issue of fitting FCU4Health into primary care. The primary themes concerned fit with the clinic’s mission and needs, clinic staffing, and patient characteristics. To this end, we first needed a new screening process with the shift to prevention of obesity and excess weight gain necessitated. Although the CORD RFA dictated that the pediatrician was to use the child’s BMI to initiate referral to FCU4Health after providing counseling, consistent with Healthcare Effectiveness Data and Information Set (HEDIS) procedures for children with BMI ≥ 85th percentile for age and gender, our pilot trial experience and CAB work group on integration confirmed that these steps and personnel aligned with clinic practices and were preferred. This modification was a substitute for the FCU screening for child problem behaviors (e.g., oppositionality) and family risks for ineffective parenting (e.g., parental depression). Next, we shortened/condensed the number of contacts between the family and the FCU4Health coordinator for the “check-up” portion of the program based on pilot data indicating that families had difficulty completing the typical 3 sessions of the FCU and preferred fewer contacts (63). FCU4Health combines the initial interview and assessment², whereas these were originally separate meetings in FCU.

Modifications to the family assessment, which were mostly additions, were fairly extensive. In the original FCU, an ecological family assessment is conducted to gather information on the various influences on both child problem behaviors and on parenting effectiveness (24). The majority of the constructs and items in the original assessment were retained because they are relevant to health behaviors (e.g., child self-regulation) or to parenting (e.g., social support, parental depressive symptoms). However, a health module was added to the survey portion of the assessment to gather more pertinent information about the constructs of (a) child dietary habits; (b) family health routines (mealtimes, sleep, media) and behaviors (dietary practices, exercise); (c) health-related quality of life; (d) weight-related stigma; (e) body image; and (f) the management of common co-occurring health conditions when present (i.e., asthma, diabetes). These additional constructs were in part a result of a working group meeting of our CAB on components of the FCU4Health. In this meeting weight-related stigma came up as an area of particular importance, as did the need for referral resources to support child and family health in the areas listed previously.

The FCU assessment also includes an observational component to rate parenting skills and family functioning using the Family Interaction Task (FIT), which is a series of semi-structured family interactions that are coded using a validated system (74). In the spirit of shortening the FCU4Health, we modified the number, length, and prompts of the FIT. In FCU, five tasks (5 min each) were administered with a focus on factors related to preventing child problem behaviors (e.g., monitoring the child’s whereabouts and peer network). In FCU4Health, we administer three tasks (4 min each) concerning health goals and promoting healthy behaviors. For example, the instructions for the Goal Setting task are:

To Child: I’d like you to talk about your goals for yourself for exercise and your diet, especially developing healthy habits. Then, please talk about how you feel that it is going right now.

To Caregiver(s): When (child name) is finished, please talk about your goals for his/her health, diet, and exercise behavior. Share with (child name) some specific ways you plan to help support those goals. Then please talk about your hopes and plans for your son’s/daughter’s future health.

The same coding system is used to assess parenting skills with one salient addition: parents’ knowledge of national guidelines for children’s health behaviors. Examples include the current recommended amount of daily physical activity, servings of fruits and vegetables, and amount of screen time. The FCU4Health developers piloted a version of these FIT prompts in the pilot trial and they were refined with guidance from our partner agency staff and the CAB.

A final addition to the family health routines assessment for FCU4Health was anthropometric evaluation of the child and the family members using a portable medical-grade electronic scale to obtain weight and body composition data. In the grant application, we proposed to capture these data from the child only. There was a concern with respect to our economic assessment as to whether we would see cost-benefit within 1 year.

Because adult BMI is more proximally linked with expensive health outcomes, and because we theorized that by promoting healthy diet and physical activity in the family, parents may also experience reductions in obesity. The CAB felt that having the entire family get on the scale would normalize the measurement process for the child. Consequently, we decided to add parent weight and body composition to the assessment and encourage being weighed and measured. As weight was not a target of the original FCU, this element was simply added³.

The purpose of the assessment is to identify services that would help families support child health and motivate parents to engage in those services. These follow-up services can take one of two forms. To address needs related specifically to parenting, the coordinator provides parenting skills training using Every day Parenting (75), a 12-module skills-based curriculum focusing on three core areas of parenting and family management: relationship quality, positive behavior support, and monitoring

---

²In the Raising Healthy Children study, we conduct the assessment first in a stand-alone contact in order to double-blind the collection of baseline data prior to randomization and follow-up assessments at later waves. The initial interview is then combined with the feedback. In routine implementation of FCU4Health, the initial interview and assessment would be combined.

³In a previous trial of FCU, children’s height and weight data was collected as part of the research protocol, but this information was not used in the intervention in any way [see Montaño et al. (42)]. Thus, we consider its explicit use in the FCU4Health an addition.
and limit setting. This element was refined from the FCU, which also shares this explicit goal, by using examples that specifically focus on health behaviors (e.g., setting limits on screen time). In both programs, the number of Everyday Parenting modules and the type and number of referrals for community-based support services are individualized to the specific needs of each family following a feedback session that discusses the key findings of the family assessment. Although FCU4Health adheres to the Everyday Parenting modules as they pertain to skill-building, because of the program’s target, coordinators add a focus on children’s nutrition and age-appropriate health behavior expectations. This element aligns with our added category in the FIT assessment of parent understanding of health guidelines. FCU does not provide this information as part of standard protocol.

To address other areas of need, the coordinator shares information about resources in the community and provides motivational and logistical support to families to connect with those resources. In FCU4Health, there is an emphasis on referrals to health-related community supports, such as food banks, community gardens, and recreational programs, and also to social services that can help the family address social determinants of health related to childhood obesity (76). There is an explicit goal of assisting families in procuring insurance for their child(ren) and securing or maintaining employment. Similar to the original FCU, FCU4Health also commonly refers parents to specialty mental health services, when indicated, for such issues as children’s mental health concerns (e.g., developmental delays, attention deficit hyperactivity disorder) parental depression and substance use. FCU4Health has a greater focus on specialty health care for common co-occurring conditions, most notably chronic health conditions such as asthma and diabetes. The CAB workgroup meeting on program components yielded recommendations on how to compile, maintain, and disseminate up-to-date information on referral resources to facilitate referrals.

Training and Evaluation Modifications

The FCU4Health training and supervision process and implementation monitoring system remains largely consistent with the most recent trials of FCU. Three aspects differ from prior trials of FCU. First, given that the Raising Healthy Children project is an effectiveness trial, the amount of consultation from FCU4Health developers and supervisors, and ongoing oversight more generally, is less prescribed in amount and duration compared to efficacy trials by Dishion and colleagues [e.g., Dishion et al. (24)]. The best comparison to our procedure and amount of training and consultation is an effectiveness-implementation hybrid type I trial conducted with the original FCU in community mental health agencies (54). A second modification to training compared to previously published research trials is the use of an e-learning course developed by Dishion and colleagues at the ASU REACH Institute as the prerequisite to in-person training in the program. The e-learning course is on the original FCU (77) and the Everyday Parenting Curriculum (78) and covers the theoretical background and core components of the parenting aspects of the program; we focused on supplementing this information with the health-related adaptations of the FCU4Health during the in-person training. Third, FCU uses a validated, observational coding system called the COACH (79) to monitor delivery of the program. The COACH is an observational rating system of fidelity in delivering FCU4Health. Skills in five areas (Conceptually accurate to FCU4Health; Observant and responsive to client needs; Actively structures the session; Corrective feedback is provided; Hope and motivation) are rated on a scale from 1 (low) to 9 (high) by trained coders. Scores on the COACH have been found to be reliable and related to change in both parenting skills and child behaviors in previous trials (80–82). For Raising Healthy Children, we are using the COACH, but are also developing an automated system to rate fidelity based on existing, validated systems for core elements of motivational interviewing (e.g., presence of complex reflections, open-ended questions) (83–85) and other family-based and parent training interventions (86, Li et al., submitted). This system will allow us to evaluate fidelity to FCU4Health for every session rather than the typical practice of coding a small sample due to the burden of observational assessment. Thus, the training of FCU4Health includes: for coordinators, completion of a 7-module e-learning course, a 3-day in-person training, completion of a mock case, and close supervision for the first two families seen is encouraged but not required, and varies based on ratings of fidelity to the protocol; for interviewers (those completing the assessments), a 1-day training that includes a practice administration; and for referring physicians and other healthcare staff and leadership/managers in the clinics, a 30–45 min orientation to the program and the referral procedures and inclusion criteria.

Specifying the Behavior Change Techniques of the FCU4Health

JaKa et al. (27) developed a standardized protocol to specify the type and amount of behavior change techniques used in behavioral interventions for pediatric obesity. They drew from the original 93 techniques in the Behavior Change Taxonomy (87). For the purposes of this article, we specify the type and rate the emphasis given each technique on a scale of 1 (low) to 5 (high). JaKa et al. also code the amount of each technique, but this can only be determined from observation of the program’s delivery. Figure 3 provides a side-by-side comparison of the FCU4Health and the original FCU. Because the FCU4Health is individually tailored to the needs of each child and family, we further specify whether a given technique is universal (received by all families in the program) or applied selectively based on needs identified in the family health behaviors assessment. We present the 23 techniques JaKa et al. (27) found to be reported at least once in their evaluation of intervention protocols, manuscripts, and workbooks, indicating salience for childhood obesity programs whereas the remaining 70 techniques are unlikely to be relevant. Further, given that FCU4Health is a family-based intervention, certain behavior change techniques are taught to caregivers to then use with the child. For example, we train caregivers to provide effective social rewards to the child to reinforce desired behaviors.
DISCUSSION

This paper presents the adaptation of an evidence-based prevention program to scale-out to a new delivery context and for a new clinical target. The well-established FCU program was adapted for the pediatric primary care context and enhanced to more effectively prevent obesity and excess weight gain in children. The resulting program, FCU4Health, is likely to be acceptable and feasible based on pilot study data (63) and is currently being tested in a large randomized effectiveness-implementation hybrid trial (26) to gather this data alongside evidence of clinic impact on children’s weight and health behaviors.

A number of key changes were made when developing FCU4Health, while other core components and characteristics of FCU were retained. Context adaptations, most of which were made a priori by the program developers, included the identification and referral process; a shift to a health promotion focus; reducing the number of total contacts for the “check-up” component; and a division of responsibilities among clinic staff for the various components of the program. Important context characteristics that remain unchanged from the original FCU are an emphasis on underserved children disproportionately at risk for the target outcome; the coordinator’s being behavioral health professionals; 1-on-1 delivery format to maximize flexibility of delivery and reduction of barriers to maximize participation; and engaging caregivers in multiple “check-ups” to track progress and continually enhance motivation to change behaviors. Content adaptations were almost exclusively due to the change in clinical target. These additions began during pilot testing of the FCU4Health and were later refined in collaboration with the CAB. A key modification was making the family assessment more relevant to weight-related behaviors. Importantly, the critical processes that underlie change in the FCU were retained. These include assessment, feedback, motivation enhancement, coordination with community supports and programs, and individualized intervention planning. Training, supervision, and implementation monitoring largely remains unchanged, with the exception of using a recently-completed e-learning course to train coordinators and developing an automated fidelity coding system as part of the ongoing Raising Healthy Children study.

Adaptation is a commonly used implementation strategy to better align EBPs with the characteristics of real-world service delivery systems and the populations being served (18). Despite the prevailing use of EBP adaptation in practice, this paper provides a number of unique contributions to the

<table>
<thead>
<tr>
<th>Behavior change technique</th>
<th>FCU</th>
<th>FCU4Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Goal setting, behavior</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>1.2 Problem solving (barriers)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Problem solving (solutions)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Action planning</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1.5 Review goals (review progress)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Review goals (consider revising)</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1.9 Commitment</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2.3 Self-monitoring of behavior</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3.1 Social support, unspecified</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Social support, unspecified (in session)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3.2 Social support, practical</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4.1 Instruction on how to perform a behavior</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5.1 Information about health consequences</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7.1 Prompts and cues</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>8.1 Behavioral practice rehearsal</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>8.3 Habit formation</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10.1 Material incentive</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>10.4 Social reward</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10.6 Non-specific incentive</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10.7 Self-incentive</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>12.1 Restructuring the physical environment</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>12.5 Adding objects to the environment</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>13.1 Identification of self as a role model</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>15.3 Focus on past success</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>99.01 Information gathering</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Notes: □ All families. □ Selected families. Rating scale: 1 = low, 5 = high emphasis.

FIGURE 3 | Specification of behavior change techniques in FCU4Health and the original FCU. All families. Selected families. Rating scale: 1, low; 5, high emphasis.
implementation research literature, such as how adaptations are characterized, for what purposes, and how the decisions to adapt were made. First, we framed our adaptation process and aims on the new implementation science concept of scaling out (15). In contrast to the common approach of incremental adaptation typical in the existing literature, where modification to either the delivery context or to the population is described, this paper is an example of simultaneous adaptation to both dimensions, which speeds translation of EBPs. We also delved deeper into the scaling out concept in an important way by providing detailed, hierarchical levels of evidence to be applied when scaling out involves adapting a program for a new clinical target. By providing Minimal, Preferred, Preferred Plus, and Optimal levels of evidence, researchers, reviewers, and stakeholders can better evaluate the case for scaling out in this manner. In combination with available support for changing delivery context or population (e.g., age, racial/ethnic group), the level of evidence can be used to justify scaling out to a new clinical target. Future work could involve similarly specifying levels of evidence for changing to a new delivery context or population characteristic (without changing the clinical target). Currently, this does not exist explicitly.

Second, we used the Stirman et al. (25) adaptation coding system in a novel way to characterize the types of adaptations made, by whom, and based on what sources of data. Stirman et al.’s system provides a framework for a comprehensive description of adaptations. We found it to be particularly useful for the current paper as it was intended to be applied retrospectively. In contrast to the typical use of the system for coding individual sessions, we were able to successfully apply it to the program as a whole using a common language that other adapters could also use to describe their adapted EBPs. The consistent use of terminology is a critical challenge in implementation research (88).

Third, we provide a comparison of the behavior change techniques used and their levels of emphasis and application between the original FCU and the new FCU4Health program. We used the techniques that JaKa et al. (27) identified as most common in EBP protocols of behavioral interventions for pediatric obesity. Specification of techniques in this manner helps to open the “black box” of how these interventions work and allows for comparison with the active ingredients of other, similar programs. In this paper, it was also useful in highlighting the similarities and the differences between FCU and FCU4Health. The differences were minor and centered on a greater emphasis in FCU4Health on changing the physical environment and the caregiver being a role model for healthier child behaviors. These minor differences provide support for our assertion that the core components responsible for the effectiveness of the original FCU were retained in FCU4Health. Last, without the ability to rate how frequently each technique was used in FCU4Health delivery, as the JaKa et al. rating system was intended, we used an emphasis scale and indicators of either to all families or to select families to further illustrate the degree of likely use. Observational coding of FCU4Health sessions in the future could be done to quantify with better precision the frequency at which each technique is used.

CONSIDERATIONS

One of the challenges in both adapting the FCU and in describing it in this paper is that the program is individually-tailored and delivery is flexible by design. This made it challenging to code adaptations; many of the elements of FCU4Health would be acceptable if done within the context of the original FCU. For example, discussing a need for more physical activity and less screen time in FCU would be appropriate if it related to a concern raised by the caregivers even if it’s not an explicit target of that program. FCU4Health more or less uses the core intervention techniques and process of the FCU, but shifts the focus to a new clinical target, pediatric obesity, and emphasizes parental management and supports to improve child health behaviors. While many adaptations described here could be considered fidelity-congruent within FCU, but not necessarily prescribed, they should be considered necessary for high fidelity to FCU4Health given the new clinical focus.

From a practical perspective, the elements that we added to FCU4Health’s questionnaires increased the time to complete, particularly for children with additional chronic health conditions (the presence of asthma and diabetes trigger additional questions). The family assessment is already a challenge to complete in many service delivery systems due to the time required. Thus, we expect in the future to pare down the assessment to its necessary constructs based on the findings of this study to reduce burden on families and agencies. This consideration harkens back to the framework of scaling out to a new delivery context and the need to consider capacity and readiness to adopt and deliver an EBP. Although assessment is commonplace in pediatric primary care, the measures are typically screeners that are very short. Moreover, assessments are sometimes administered via semi-structured interview format where pediatricians write out responses. Thus, it might be challenging to change this practice in favor of the FCU4Health questionnaire and FIT assessment even if the time required is comparable.

CONCLUSIONS

This paper provides a detailed account of the many sources of information and data that inform an ongoing process of adaptation to meet the changing needs of the setting and the population served. Our approach aligns well with the Dynamic Sustainability Framework (13) and is an example of community engaged dissemination and implementation (65). Our process to date occurred over 6 years and will continue. Adaptations have and will continue to occur as we triangulate data from multiple sources (e.g., delivery, feedback from stakeholders, examination of clinical effects). As the FCU4Health is implemented, we are continuing to refine the program components and delivery strategies with input from our CAB, the partner agencies, FCU4Health coordinators, and our implementation support staff. A key activity as the Raising Healthy Children study nears completion is to review our implementation data and work with stakeholders, including caregivers and children, to determine what the program will look like and how it will be delivered.
in the “real world”; that is, as the agencies attempt to sustain implementation of FCU4Health outside of a formal research study. We expect the process of adaptation to be ongoing as the healthcare landscape for children evolves and the priorities of the agencies that serve them also shift.

**AUTHOR CONTRIBUTIONS**

JS and CB conceived of the study. JS, CB, JR-S, ZM, AM, AC, MB, and TD participated in different phases of the adaptation process and contributed to the adaptation of the FCU in substantive ways. JS, CB, and JR-S collaborated in the writing of the manuscript and wrote the final manuscript. All authors have read and approved the manuscript.

**FUNDING**

This study is supported by grant U18 DP006255 from the National Center for Chronic Disease Prevention and Health Promotion of the Centers of Disease Control and Prevention, under the Childhood Obesity Research Demonstration Project 2.0 (CORD), awarded to CB and JS. Additional support was provided by grant P30 DA027828 from the National Institute on Drug Abuse, awarded to C. Hendricks Brown; minority fellowship SM60563-40 awarded to ZM and minority fellowship SM060563-03 awarded to AC from the Department of Health and Human Services; and the Implementation Research Institute (IRI) at the George Warren Brown School of Social Work, Washington University in St. Louis through grant R25 MH080916 from the National Institute of Mental Health and the Department of Veterans Affairs, Health Services Research and Development Service, Quality Enhancement Research Initiative (QUERI). The opinions expressed herein are the views of the authors and do not necessarily reflect the official policy of position of the Centers for Disease Control and Prevention, the Department of Veterans Affairs, the National Institute on Drug Abuse, the National Institute of Mental Health, or any other part of the US Department of Health and Human Services. Development work for this project was supported by a research grant from the College of Liberal Arts and Sciences at Arizona State University, awarded to TD.

**ACKNOWLEDGMENTS**

The authors wish to thank our collaborators on this cooperative agreement in the Obesity Prevention and Control Branch of the Division of Nutrition, Physical Activity, and Obesity at the Centers for Disease Control and Prevention; our dedicated project staff at Arizona State University, Northwestern University, the University of Southern California, and the University of Washington; our partner healthcare agencies; and the many individuals providing guidance and input as members of our community advisory board, as well as the families that have participated in the many activities mentioned in this article.

**REFERENCES**


**Conflict of Interest Statement:** JS and CB led the adaptation of and co-developed the Family Check-Up 4 Health program along with TD. TD is the developer of the original Family Check-Up program.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Smith, Berkel, Rudo-Stern, Montaño, St. George, Prado, Mauricio, Chiapa, Bruins and Dishion. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Using Instructional Design, Analyze, Design, Develop, Implement, and Evaluate, to Develop e-Learning Modules to Disseminate Supported Employment for Community Behavioral Health Treatment Programs in New York State

Sapana R. Patel1,2*, Paul J. Margolies1,2, Nancy H. Covell1,2, Cristine Lipscomb3 and Lisa B. Dixon1,2

1 The New York State Psychiatric Institute, Research Foundation for Mental Hygiene, New York, NY, United States, 2 Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY, United States, 3 Intrac Inc., Instructional Design and Learning Strategy, Reno, NV, United States

Background: Implementation science lacks a systematic approach to the development of learning strategies for online training in evidence-based practices (EBPs) that takes the context of real-world practice into account. The field of instructional design offers ecologically valid and systematic processes to develop learning strategies for workforce development and performance support.

Objective: This report describes the application of an instructional design framework—Analyze, Design, Develop, Implement, and Evaluate (ADDIE) model—in the development and evaluation of e-learning modules as one strategy among a multifaceted approach to the implementation of individual placement and support (IPS), a model of supported employment for community behavioral health treatment programs, in New York State.

Methods: We applied quantitative and qualitative methods to develop and evaluate three IPS e-learning modules. Throughout the ADDIE process, we conducted formative and summative evaluations and identified determinants of implementation using the Consolidated Framework for Implementation Research (CFIR). Formative evaluations consisted of qualitative feedback received from recipients and providers during early pilot work. The summative evaluation consisted of levels 1 and 2 (reaction to the training, self-reported knowledge, and practice change) quantitative and qualitative data and was guided by the Kirkpatrick model for training evaluation.

Results: Formative evaluation with key stakeholders identified a range of learning needs that informed the development of a pilot training program in IPS. Feedback on this pilot training program informed the design document of three e-learning modules on IPS: Introduction to IPS, IPS Job development, and Using the IPS Employment Resource.
Background and Rationale for Educational Activity

A recent report by the Institute of Medicine, *Best Care at a Lower Cost: The Path to Continuously Learning Health Care in America* (1), reported, “Achieving higher quality care at lower cost will require fundamental commitments to the incentives, culture, and leadership that foster continuous learning, as the lessons from research and each care experience are systematically captured, assessed, and translated into reliable care.” Central to the translation from research to practice and reliable care is training health-care providers in evidence-based practices (EBPs). In behavioral health care, training in EBPs often involves developing new clinical competencies. This training should take into account the context and needs of the practice community as well as strategies to facilitate adoption and implementation (2). Increasingly, training utilizes online modalities to expand its reach and efficiency, digital media to promote active engagement, shorter learning sessions to foster knowledge retention, and methods to demonstrate and practice skills that can be applied in the workplace (3).

Implementation science, a field dedicated to understanding targeted dissemination and implementation of EBPs and the use of strategies to improve adoption in community health-care settings, has guided the work of translating research to practice. Numerous frameworks in implementation science have been associated with effective implementation. Damschroder (Figure 1) (4) combined 19 published implementation theories into the Consolidated Framework for Implementation Research (CFIR). The CFIR provides a menu of constructs that have been associated with effective implementation. The framework is organized into five domains: intervention characteristics, inner setting, outer setting, characteristics of individuals, and process. Under the inner setting domain, one key construct under the component readiness for implementation is access to knowledge and information. Access to knowledge and information is defined as the ease of access to digestible information and knowledge about the practice and how to incorporate it into work tasks. This is the function of training. It is purported that when timely on-the-job training is available, implementation is more likely to be successful (5).

Although training is an important determinant of successful implementation (5), the field of implementation science lacks systematic approaches for the development of training that takes into account the learners’ needs, context, and optimal modalities for learning. Training in EBPs and their evaluation has been identified as a priority item on the National Institutes of Health research agenda (Program Announcement in Dissemination and Implementation Research in Health; https://grants.nih.gov/grants/guide/pa-files/PAR-16-238.html) and a commonly used implementation strategy in implementation practice and research (6). Intervention or EBP developers are not likely to have expertise in instructional design and may miss the mark of engaging busy practitioners in training for several reasons. First, didactic approaches may not take into account the level of interest or needs of the practitioners. Second, traditional training approaches may not consider organizational factors (i.e., time available for provider training) also key to successful implementation (7). Third, how individuals learn and process information is evolving given our access to the Internet and technology. The field of implementation science may benefit from an ecologically valid approach to the development of learning experiences for training health-care practitioners.

Recent reports have pointed to the utility of instructional design in the dissemination and implementation of EBPs in behavioral health (8, 9). The field of instructional design offers one model, Analyze, Design, Develop, Implement and Evaluate (ADDIE) (10), that takes into account learning theory, the learner’s needs and environment, and approaches to training practitioners in EBPs. The foundations of ADDIE are traced back to World War II when the U.S. military developed strategies for rapidly training people to perform complex technical tasks. The ADDIE model is used in creating a teaching curriculum or a training that is
geared toward producing specific learning outcomes and behavioral changes. It provides a systematic approach to the analysis of learning needs, the design and development of a curriculum, and the implementation and initial evaluation of a training program (11, 12). This model of developing training programs is particularly useful if the focus of the program is targeted toward changing participant behavior and improving performance. ADDIE is increasingly being adopted in industries such as health care (13). Recent studies have successfully adopted the ADDIE model to improve patient safety, procedural competency, and disaster simulation (14, 15). It has also been effectively used in medical training and education to change practice behaviors in the management of various medical conditions (16–18).

The options for delivery and modalities used in training (e.g., mobile devices, webcasts, and podcasts) have expanded significantly in the last decade. With online learning technology, there is an opportunity to reach learners anytime and anywhere to provide performance support. One such example of an online learning technology is e-learning modules. e-Learning modules are self-paced lessons that enable the learner to read text, listen to narrated content, observe video scenarios, and respond to questions or prompts, in a multimedia format designed to maximize engagement and retention. Learning management systems (LMSs) host e-learning modules and capture learning metrics and performance. Learning analytics provided by LMSs enable the ability to track individual and group performance which may be used to provide feedback and support continuous learning in large systems of care.

In this report, we provide an example of the application of ADDIE in the development and evaluation of e-learning modules as one strategy among a multifaceted approach to the dissemination and implementation of the individual placement and support (IPS) model of supported employment in community treatment programs in New York State (NYS). Specifically, we (1) describe the application of an instructional design framework, ADDIE, in the iterative development of e-learning modules for IPS; (2) conduct a large-scale dissemination of the IPS e-learning modules throughout the state using an LMS; (3) evaluate learner reaction, self-reported knowledge, and practice change after IPS e-learning modules; (4) identify key barriers and facilitators to future IPS implementation using formative and summative ADDIE evaluation data and the CFIR.

### PEDAGOGICAL FRAMEWORKS

We used three frameworks to guide the process of developing e-learning modules (ADDIE), identify determinants of future training and implementation (CFIR), and evaluate the IPS e-learning modules [Kirkpatrick model (19)]. The ADDIE model consists of five phases, beginning with identifying key stakeholder needs, educational goals, and optimal methods of content delivery (analysis). This information was used to establish a design

#### FIGURE 1 | Using an instructional design framework [Analyze, Design, Develop, Implement, and Evaluate (ADDIE)] as a systematic process to develop training in an implementation framework [Consolidated Framework for Implementation Research (CFIR)].
document for the training (design) that is vetted by key stakeholders prior to building the e-learning modules (development). After iterative refinement, e-learning modules were disseminated and evaluated using the Kirkpatrick model for training evaluation (19) (implementation/evaluation). Results from the formative and summative evaluations conducted during the ADDIE process, identified barriers/facilitators to implementation using CFIR domains (Figure 1). Doing so allowed the IPS team to iteratively ensure sufficient attention to contextual variables, align with the larger conceptual and empirical implementation literature (9, 20) as well as select strategies to build a multifaceted approach to IPS implementation.

LEARNING ENVIRONMENT

In November 2007, the NYS Office of Mental Health (OMH) and the Department of Psychiatry, Columbia University, established the Center for Practice Innovations (CPI) at Columbia Psychiatry and New York State Psychiatric Institute to promote the widespread use of EBPs throughout NYS. CPI uses innovative approaches to build stakeholder collaborations, develop and maintain providers’ expertise, and build agency infrastructures that support implementing and sustaining these EBPs. CPI works with OMH to identify and involve consumer, family, provider, and scientific-academic organizations as partners in supporting the goals of OMH and CPI. CPI’s initial charge was to provide training for the NYS behavioral health-care workforce. Given the size and geographical dispersion of this workforce, CPI turned to distance-learning technologies and e-learning modules (21, 22). Distance technologies may offer cost-effective alternatives to typical training methods, and some evidence suggests that such technologies are at least as effective as a face-to-face training (21). CPI has collaborated with key stakeholders and content experts to create more than 100 e-learning modules to provide training for its initiatives. CPI’s online modules and resources require the use of an online learning platform, an LMS, that facilitates access to online training, event registration, and resource libraries for each initiative.

One of these initiatives, IPS, provides training and implementation support in an evidence-based approach to supported employment (23). Rates of competitive employment were low across NYS, with a competitive employment rate of 9.2% in 2011 prior to systematic IPS implementation (Patient Characteristics Survey data, 2011 obtained from https://www.omh.ny.gov/omhweb/statistics/). In response, OMH leadership identified supported employment as a key service in personalized recovery oriented services (PROS) programs, a comprehensive model that integrates rehabilitation, treatment, and support services for people with serious mental illness. The number of PROS programs in New York has increased significantly over the past decade: in 2017, 86 programs were serving 10,500 individuals. In order to reach these 86 programs statewide, the IPS initiative developed a series of three e-learning modules: Introduction to IPS, IPS Job Development, and Using the IPS Employment Resource Book. The module development team included an instructional designer, subject matter experts (SMEs), course developers, and a project manager.

PEDAGOGICAL FORMAT: E-LEARNING MODULE DESIGN USING ADDIE

Analysis: Learning Objectives

In the analysis phase, the instructional problem was clarified, the instructional goals and objectives were established, and the learner’s environment, existing knowledge, and skills were identified. The module development team engaged in a discussion to identify the instructional problem and understand the expectations for performance after completing the modules. Because IPS had not been previously implemented in NYS, it was expected that learners’ existing knowledge and skills of IPS would be minimal. Formative evaluation via preliminary discussions with agency administrators, employment supervisors, and employment staff members in PROS programs included questions about learners’ experiences with and opinions about traditional vocational rehabilitation methods, attitudes about IPS principles (i.e., zero exclusion), awareness of or experiences with IPS, and expectations and attitudes about the likelihood of program recipients in their programs working competitively. These discussions revealed several needs: lack of understanding of the evidence for IPS (CFIR: intervention), discomfort with some IPS principles which are inconsistent with traditional approaches to vocational rehabilitation (characteristics of individuals), lack of knowledge about the specific skills and tasks involved in the model (characteristics of individuals), lack of familiarity with how to do job development and why it is important (characteristics of individuals), and the lack of tools that can be used in real-time meetings with potential employers (implementation process). These data informed the development of a curriculum for a pilot training program in IPS that consisted of in-person training, webinars, and on-site technical assistance. Through this pilot process, observations were made about learners’ strengths and additional training needs, and the PROS program environment. In addition, program recipients’ (adults diagnosed with serious mental illness, living in the community, many with histories of hospitalizations and treatment) employment needs (e.g., consistent with individuals’ personal strengths and interests), part-time for many, easily accessible with public transportation (outer setting) supported another cycle of modifications to the IPS curriculum and informed decisions about pedagogical format. As the initiative required scalability across the state of New York, it was determined that e-learning modules would be an important resource-efficient implementation strategy.

Design

The design phase established learning objectives, exercises, content, lesson planning, and media selection via a design document, which served as the blueprint for building the training program. The instructional designer gathered feedback from the analysis phase and resources on the topic provided by SMEs (e.g., books, research publications, information available online) and identified content to support the learning objectives (Table 1) for all three IPS e-learning modules. The module development team designed a 10-item knowledge quiz and a 10-item level 1 reaction survey.
TABLE 1 | Learning objectives for individual placement and support (IPS) modules.

| Introduction to IPS | 1. Learn about the historical context for IPS  
| | 2. Learn about why employment is important  
| | 3. Learn core practitioner skills that follow fundamentals of IPS and its implementation |
| IPS Job Development Module | 1. Understand the importance of job development and the employment specialist role  
| | 2. Learn about the role of the treatment team and how each member can support job development  
| | 3. Learn about how to coach consumers as they meet employers, build networks, discuss disclosure |
| Using the IPS Employment Resource Book | 1. Learn about the Employment Resource Book and its usefulness for consumers, family members, and providers |

consisting of both closed- and open-ended questions. Iteratively, the instructional designer presented design documents for review and feedback from the module development team. An example design document for the IPS Job Development module is provided in the Figure S1 in Supplementary Material.

Development

During the development phase, the course developer received the reviewed design document and used an authoring tool software to create multimedia e-learning modules according to the design document. During this phase, the IPS modules were animated using video, graphics with narration, knowledge checks, and photographs. Formative evaluation from the analysis phase led to the development of a tool, the Employment Resource Book (24), that could be utilized by key stakeholders (providers, supervisors, and recipients) during any phase of employment (e.g., considering work, actively seeking employment, maintaining employment), and one module was developed to provide guidance about using this resource. The IPS training was built into three short e-learning modules to reflect learners’ time availability and attention span during the workday, then tested in prototype with the module development team and revised.

Implementation

During the implementation phase, e-learning modules were uploaded to the CPI LMS for usability testing. During usability testing, the module's functionality is evaluated prior to training implementation. For example, the module development team tested whether videos play and navigation works (e.g., next buttons and links to additional resources) on a variety of web browsers and devices. Feedback from the usability testing phase is used to fix errors in navigation and improve user experience (25). After usability testing issues were addressed, the modules were ready for implementation.

When the IPS initiative began, the NYS-OMH Rehabilitation Services Unit sent an official email communication strongly encouraging PROS program providers and supervisors to participate in the training offered by the CPI IPS initiative. Further, each PROS program supervisor received an email, alerting them that the new IPS e-learning module was available in CPI's LMS. Through the LMS, PROS program participation in the modules was tracked, and completion could be monitored by PROS programs and NYS-OMH.

Evaluation

We applied quantitative and qualitative methods as part of formative and summative evaluation in the ADDIE process. Formative evaluations consisted of qualitative feedback received from recipients and providers during early pilot work, which identified training needs. The summative evaluation consisted of quantitative and qualitative data and was guided by the Kirkpatrick model for training evaluation (19). The four levels of evaluation are (1) the reaction of the learner about the training experience, (2) the learner's resulting learning and increase in knowledge from the training experience, (3) the learner's behavioral change and improvement after applying the skills on the job, and (4) the results or effects that the learner's performance has on care provided. For this report, we focus on the first two levels, specifically, the level 1—reaction of the learner including training experience, self-reported knowledge acquisition, and self-reported practice change through a survey and level 2—resulting knowledge through post-module quizzes.

To keep the learner experience seamless, a decision was made to embed the knowledge quiz, assessing knowledge of IPS model-related concepts, skills, and tools, within each module. In order for the module to be marked as completed, learners are required to answer at least 80% of the knowledge items correctly, which satisfies continuing education accreditation requirements. Learners are able to retake the quiz as many times as needed to meet this criterion score. Once the module is completed, the learner is prompted to complete the level 1 survey. The level 1 reaction survey was based on learning objectives set forth in each e-learning module, accreditation requirements, and example questions from Kirkpatrick level 1 (19). Questions included rating the module overall, if it met stated learning objectives, if the information presented was new to the learner, and questions about module-specific self-reported knowledge and practice change. In addition, three open-ended questions were included: What could we improve? What do you like the most about this module? and Where do you think you might use what you learned in this module?

The NYS Psychiatric Institute Institutional Review Board determined that this evaluation did not meet the definition of human subject research.

Analysis

Using IBM® SPSS® Statistics Version 24, we applied descriptive statistics to quantitative level 1 summative evaluation data. For the qualitative formative and summative evaluation data, we employed a thematic analysis to identify themes within the open-ended question data (26). Two coders reviewed the open-ended question data independently to identify codes and develop an initial code list. The coders combined codes into overarching themes and met to review and label them. Coders met twice to discuss discrepancies and achieve consensus on key barriers and facilitators within the CFIR framework. We report on those themes that were raised by at least 10% of the sample.
RESULTS

We describe the inputs and outputs during each phase of IPS module development using ADDIE in Table 2. Formative evaluation during each stage of ADDIE allowed for the iterative revision of the content for each module and the identification of needs for subsequent modules. Feedback received from the evaluation of the first module led to the development of the second module (i.e., desire to learn more about job development) and to the development of the Employment Resource Book including the associated third module (i.e., desire to be better equipped to deal with common challenges).

Summative evaluation examined the impact of the IPS training modules and assisted in the identification of barriers and facilitators for IPS implementation in the future. Table 3 summarizes level 1 evaluation data for all three IPS modules. Learners’ background and experience varied considerably across programs. Many were rehabilitation counselors, social workers, and some had non-behavioral health backgrounds. Learners rated all three modules highly (mean range: 4.4–4.5 out of 5). Learners also indicated that the modules presented new information and met their stated learning objectives (mean range: 4.3–4.4 out of 5). Similarly, learners’ self-report of knowledge acquisition was high (mean range: 4.4–4.6 out of 5). About half of learners indicated that they would change their practice after watching the modules (range: 48–51%). All learners who completed the level 1 evaluation demonstrated 80% or better mastery of knowledge on the level 2 evaluation embedded in each module.

### Table 2 | Using Analyze, Design, Develop, Implement, and Evaluate (ADDIE) to develop individual placement and support modules.

<table>
<thead>
<tr>
<th>ADDIE model</th>
<th>Inputs</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis</td>
<td>Preliminary discussions with personalized recovery oriented service program providers, recipients, Office of Mental Health staff</td>
<td>Learning needs data Pilot curriculum and training program Revised learning objectives and format to scale training</td>
</tr>
<tr>
<td>Design</td>
<td>Learning objectives, development of evaluation, and media selection SME and module development team of design document</td>
<td>Design document Revised design document</td>
</tr>
<tr>
<td>Development</td>
<td>Authoring tool applied to design document to develop multimedia training Module development and learning management system team review of module</td>
<td>Multimedia e-learning module</td>
</tr>
<tr>
<td>Implementation</td>
<td>Assessment of technical fit and specifications Pilot launch: usability testing</td>
<td>Refinement of course</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Evaluation design and monthly review of results Level 1 and 2 evaluation results Barriers and facilitators to future implementation beyond e-learning</td>
<td></td>
</tr>
</tbody>
</table>

Open-ended question themes and related CFIR domains from these e-learning modules helped identify additional implementation support needs to be addressed by the multifaceted approach to implementing IPS (i.e., statewide webinars, regional online meetings focusing on special topics such as IPS fidelity and supervision, an IPS library with tools to help IPS implementation, and individualized program consultations that focus on addressing implementation challenges and enhancing provider competence). Themes from the open-ended questions for all three IPS modules are described using the CFIR in Table 4. These themes related to three CFIR domains: outer setting, inner setting, and implementation process. They provided information on how the modules were acceptable, what the future learning needs are, and how the information learned will be used in everyday practice.

DISCUSSION

This report provides one example of how an instructional design approach may be applied to the development of e-learning modules as one strategy in a multifaceted approach to the implementation of IPS supported employment for community program providers in a large state public behavioral health system. Through iterative development, we applied the ADDIE model to develop a series of e-learning modules for IPS. Using an LMS, these modules were disseminated and evaluated by PROS program providers throughout NY state. Results from both level 1 and level 2 evaluations indicate that the ADDIE model was successful in improving practitioner knowledge. In addition, learners received the e-learning modules favorably, rating them highly overall and noting that they met stated learning objectives and presented new information. Throughout the development process, data from the e-learning modules were described using the CFIR to identify needs that led to additional e-learning modules as well as strategies for subsequent implementation supports through a learning collaborative statewide (27).

The ADDIE model and CFIR were used as complementary approaches in the development of e-learning resources for training providers in an EBP. Our experience in this process produced several lessons learned and recommendations for implementation researchers and practitioners. The analysis phase of the ADDIE model required assessment of multistakeholder needs and context early on in the process of developing training. We recommend taking the time to assess and include end users and recipients to shape and increase the ecological validity of the training. In addition, the use of the CFIR domains allowed us to anticipate barriers and map future implementation strategies. During the design process, the establishment of clear and measurable learning objectives was important and facilitated focus and evaluation of knowledge and skill acquisition. We recommend the <em>a priori</em> assembly of e-learning module development teams to work with the instructional designer and establish an efficient process for the review of training content and format through weekly iterative review meetings during the design and development stages. Although the ADDIE process points to the introduction of the learning platform (e.g., website, LMS) at the implementation stage, we would recommend that the team with technical expertise (i.e., in our case, the courseware developers...
TABLE 3 | Level 1 data from all three individual placement and support (IPS) modules.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Introduction to IPS (M, SD)</th>
<th>IPS Job Development (M, SD)</th>
<th>Using the IPS Employment Resource Book (M, SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would rate this training (with five stars being the best)</td>
<td>4.5 (0.74)</td>
<td>4.4 (0.77)</td>
<td>4.5 (0.73)</td>
</tr>
<tr>
<td>The online module met its stated objectives (1 = very inadequately to 5 = considerably)</td>
<td>4.4 (0.60)</td>
<td>4.3 (0.76)</td>
<td>4.4 (0.60)</td>
</tr>
<tr>
<td>The module included information that was new to me</td>
<td>4.2 (0.88)</td>
<td>4.2 (0.91)</td>
<td>4.3 (0.74)</td>
</tr>
</tbody>
</table>

**Self-reported knowledge**

**IPS: Introduction to the IPS Model of Supported Employment**
As a result of this online module, I better understand the importance of employment for persons with mental illness | 4.5 (0.80) |
As a result of this online module, I better understand the rationale for and fundamentals of IPS | 4.5 (0.76) |
As a result of this online module, I better understand core practitioner skills and how to implement IPS | 4.5 (0.73) |

**IPS: Job Development**
As a result of this online module, I better understand the importance of job development | 4.4 (0.75) |
As a result of this online module, I better understand the importance of the employment specialist role | 4.5 (0.77) |
As a result of this online module, I better understand how to support job development across the treatment team | 4.4 (0.76) |

**IPS: Using the Employment Resource Book**
As a result of this training, I better understand how to access the Employment Resource Book | 4.6 (0.60) |
As a result of this training, I better understand how to use the book for guidance and direction concerning consumers’ employment goals | 4.6 (0.61) |

**Self-reported practice change**

<table>
<thead>
<tr>
<th></th>
<th>N(%)</th>
<th>N(%)</th>
<th>N(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>This activity validated my current practice; no changes will be made</td>
<td>49%</td>
<td>49%</td>
<td>52%</td>
</tr>
<tr>
<td>Change the management and/or treatment of my patients/clients</td>
<td>31%</td>
<td>28%</td>
<td>27%</td>
</tr>
<tr>
<td>Create/revise protocols, policies, and/or procedures</td>
<td>20%</td>
<td>20%</td>
<td>21%</td>
</tr>
</tbody>
</table>

* N = 523.
* N = 312.
* N = 127.
*Likert scale: 1-strongly disagree to 5-strongly agree.

TABLE 4 | Themes and CFIR domains from level 1 survey open-ended questions for individual placement and support (IPS) modules.

<table>
<thead>
<tr>
<th>Open-ended questions</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduction to IPS</strong></td>
<td>More on common challenges (symptoms developing, substance use, disclosure, failed attempts) (outer setting)</td>
</tr>
<tr>
<td></td>
<td>How to do job development (inner setting)</td>
</tr>
<tr>
<td></td>
<td>Breakdown of three visits with potential employers (inner setting)</td>
</tr>
<tr>
<td></td>
<td>Learning to be more person-centered, person-driven, de-stigmatizing (inner setting)</td>
</tr>
<tr>
<td></td>
<td>How to do rapid job search (inner setting)</td>
</tr>
<tr>
<td><strong>IPS Job Development</strong></td>
<td>One-on-one with clients (implementation process)</td>
</tr>
<tr>
<td></td>
<td>Supervision (implementation process)</td>
</tr>
</tbody>
</table>
and LMS administrators) be introduced earlier in the process during the development stage. This is crucial to the feasibility and usability of the end product. Once implemented, we recommend a scheduled monthly review of the evaluation data that is being collected as learners participate in the e-learning modules. This information will identify any needed revisions to the training content, the need for future content development, and barriers and facilitators for future implementation.

This article reports on the development of e-learning modules that were one part of a larger implementation effort in a state system. This implementation was not a part of a rigorous research evaluation. Limitations of this report include inability to formally assess pre–post knowledge, practice and readiness for IPS implementation using validated scales based on accepted standard in the literature, variation in sample sizes for the e-learning modules precluding examination of a stable cohort of learners over time, and the inability to directly assess the specific impact of these e-learning modules on employment outcomes apart from other elements of the entire initiative. Notably, only half of the providers who completed the evaluation noted an intention to change their practice, and we did not have the capacity to assess practice change at the individual provider level at this stage of IPS implementation (level 3). However, in our subsequent work (27), program fidelity assessments using established measures demonstrated improvement over time, suggesting that level 3 provider practice change and fidelity self-assessed by program sites are shown to be associated with higher employment rates (level 4), which are sustained over time (28). Future research may focus on more rigorous evaluation of knowledge, practice change, mixed-method assessment of how the content from e-learning modules influences practice, and the essential role of care recipients in helping to design training within implementation efforts.

From adoption to sustainability, implementation science focuses on strategies to promote the systematic uptake of research findings into routine practice. Successful implementation relies on iterative, interacting activities that follow a systematic process for strategy development. In the case of training as an implementation strategy, instructional design offers a systematic and iterative process. First, it applies instructional theory to the development of training regardless of subject matter. Second, it identifies fundamental elements of the learners’ needs and real-world setting factors in addition to the EBP being implemented. Third, it creates accountability to align training content with measurable learning objectives and assesses learner knowledge and skill acquisition based on content. Lastly, it engages multimedia novel approaches in the development of educational and training resources.

Compared to more intensive approaches to training and workforce development, the development of e-learning modules informed by an instructional design approach provides implementation science the opportunity to scale and support training at the level of knowledge and skill acquisition for a range of EBPs. These modules can be used either as stand-alone or as part of blended learning activities and implementation supports as in the IPS initiative. Another example, in the case of complex, multi-component intervention or model of care, is CPI’s work with Assertive Community Treatment, where instructional design is used to develop e-learning modules as a first step in a blended learning training curriculum for practitioners in a state system (29). As such, there is increasing interest in examining the effect of an instructional design approach to training in behavioral health, especially for large systems of care.

CONCLUSION

Instructional design approaches such as ADDIE may offer implementation scientists and practitioners a flexible and systematic guideline for the development of e-learning modules as a single component or one strategy in a multifaceted approach for training practitioners in EBPs. In this way, this approach facilitates the translation between science to practice that takes into account the context of the learner and leverages technology for expanded reach, both promising approaches for workforce development and a learning health-care system (1, 30).

AUTHOR CONTRIBUTIONS

SP, CL, and LD conceived the study and conceptual framework. SP and NC managed data and analyses. SP, PM, NC, and CL contributed to writing the manuscript with feedback and supervision from LD.

ACKNOWLEDGMENTS

SP is a fellow of the Implementation Research Institute (IRI), at the George Warren Brown School of Social Work, Washington University in St. Louis, through an award from the National Institute of Mental Health (R25 MH080916) and the Department of Veterans Affairs, Health Services Research & Development Service, Quality Enhancement Research Initiative (QUERI).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fpubh.2018.00113/full#supplementary-material.

REFERENCES


FIGURE S1 | IPS job development design document.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AR and the handling Editor declared their shared affiliation.
A Pragmatic Approach to Guide Implementation Evaluation Research: Strategy Mapping for Complex Interventions

Alexis K. Huynh1*, Alison B. Hamilton1,2, Melissa M. Farmer1, Bevanne Bean-Mayberry1,2, Shannon Wiltsey Stirman3,4, Tannaz Moin1,2 and Erin P. Finley5,6

1VA Greater Los Angeles HSR&D Center for the Study of Healthcare Innovation, Implementation and Policy, Los Angeles, CA, United States, 2David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States, 3Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA, United States, 4VA Palo Alto Healthcare System, Menlo Park, CA, United States, 5South Texas Veterans Healthcare System, San Antonio, TX, United States, 6UT Health Science Center, San Antonio, TX, United States

Introduction: Greater specification of implementation strategies is a challenge for implementation science, but there is little guidance for delineating the use of multiple strategies involved in complex interventions. The Cardiovascular (CV) Toolkit project entails implementation of a toolkit designed to reduce CV risk by increasing women’s engagement in appropriate services. The CV Toolkit project follows an enhanced version of Replicating Effective Programs (REP), an evidence-based implementation strategy, to implement the CV Toolkit across four phases: pre-conditions, pre-implementation, implementation, and maintenance and evolution. Our current objective is to describe a method for mapping implementation strategies used in real time as part of the CV Toolkit project. This method supports description of the timing and content of bundled strategies and provides a structured process for developing a plan for implementation evaluation.

Methods: We conducted a process of strategy mapping to apply Proctor and colleagues’ rubric for specification of implementation strategies, constructing a matrix in which we identified each implementation strategy, its conceptual group, and the corresponding REP phase(s) in which it occurs. For each strategy, we also specified the actors involved, actions undertaken, action targets, dose of the implementation strategy, and anticipated outcome addressed. We iteratively refined the matrix with the implementation team, including use of simulation to provide initial validation.

Results: Mapping revealed patterns in the timing of implementation strategies within REP phases. Most implementation strategies involving the development of stakeholder interrelationships and training and educating stakeholders were introduced during the pre-conditions or pre-implementation phases. Strategies introduced in the maintenance and evolution phase emphasized communication, re-examination, and audit and feedback. In addition to its value for producing valid and reliable process evaluation data, mapping implementation strategies has informed development of a pragmatic blueprint for implementation and longitudinal analyses and evaluation activities.
Discussion: We update recent recommendations on specification of implementation strategies by considering the implications for multi-strategy frameworks and propose an approach for mapping the use of implementation strategies within complex, multi-level interventions, in support of rigorous evaluation. Developing pragmatic tools to aid in operationalizing the conduct of implementation and evaluation activities is essential to enacting sound implementation research.

Keywords: implementation strategies, strategy mapping, complex interventions, implementation blueprint, evaluation

BACKGROUND

With rapid growth in the field of implementation science has come increasing complexity in the way that studies are planned and executed. Evidence-based interventions to improve the quality of care are frequently multi-component, comprised of, for example, both patient- and provider-facing elements (1). Implementation efforts are often large-scale and likely to be conducted across multiple sites simultaneously, each of which may have its own unique characteristics, needs, and resources (2). There is a growing array of implementation strategies—“methods or techniques used to enhance the adoption, implementation, and sustainability of a clinical practice or program” (3)—available to address the varied needs of different sites. Correspondingly, the use of implementation strategies has become increasingly sophisticated, with a growing number of efforts using a combination of strategies to target multiple levels of an organization (e.g., providers, middle managers, and high-level administrators).

There has been an increasing call for implementation research studies to describe their use of implementation strategies with greater specificity and precision, with two primary goals: replication and evaluation (4). At its most basic, this call for greater precision in the description of implementation strategies seeks to increase our ability to identify and replicate strategies that are effective in supporting adoption, scale-up, and spread of best practices in health care (4). Precise specification of how implementation strategies are used allows for greater ability to evaluate their effectiveness, understand potential mechanisms of action, and identify areas for improvement, thereby contributing to rapid evolution of the knowledge base in implementation science (3, 4). It is well recognized that there is poor replication of clinical interventions (5), and we often see the same phenomenon in implementation studies, with initially promising strategies failing to show impact in later efforts (6–8). Consequently, many implementation studies occur as isolated events, and the opportunity to build incrementally toward a knowledge base for effective implementation is compromised.

In response to this concern, a growing literature has called for standardization in implementation reporting, encouraging use of a common language for naming and defining strategies and describing their functional components (3, 9–11). Powell and colleagues (9) have done much to support this effort by developing a compilation of 73 discrete implementation strategies through a process of expert review. Waltz and colleagues (10) proposed a taxonomy for organizing those 73 strategies into nine overarching conceptual categories reflecting their core goals and approaches (e.g., involving stakeholders, education, etc.). Proctor and colleagues (3) have offered guidelines for the specification of implementation strategies, recommending that each implementation strategy be described in terms of seven domains: the actors involved, actions undertaken, action targets, timing or temporality, dose, implementation outcomes, and theoretical justification.

The development of these rubrics for defining and specifying implementation strategies has resulted in a significant change in how implementation research is described, and the level of information available to support understanding and interpretation of findings. For example, Bunger and colleagues (11) developed a method for using activity logs as part of a multi-component effort to improve children’s access to behavioral health services. Use of these detailed logs facilitated the identification of discrete strategies enacted over time, while also supporting documentation of the implementation activities, intent, duration, and actors involved. This documentation allowed for more precise estimation of the effort involved. Gold and colleagues (12) engaged in similar description of implementation strategies operationalized as part of a diabetes quality improvement intervention occurring in commercial and community healthcare settings. They found that, while the strategies utilized and outcome observed were constant across settings, specific components of the strategies used—including actor, action, temporality, and dose—were adapted to fit local contexts, thus underscoring the importance of flexibility in implementation (12). Most recently, Boyd and colleagues (13) coded implementation team meetings to characterize implementation strategies. They identified six categories of strategies: quality management, restructuring, communication, education, planning, and financing, including one (communication) that had not been identified as such in previous taxonomies. In preliminary analyses, financing was associated with greater intervention fidelity. In another recent study, Rogal and colleagues used an electronic survey to assess use of specific strategies in implementation of evidence-based hepatitis C treatment (14). In doing so, they were able to identify 28 strategies that were significantly associated with initiation of evidence-based hepatitis C treatment, including use of data warehousing techniques and intervening with patients. Collectively, these studies have been pioneering in their use of the shared language offered by Proctor and colleagues (3) to achieve consistent reporting in implementation research; they point the way forward for future efforts.

Nonetheless, movement toward greater specification of individual implementation strategies raises challenges, particularly
related to reporting on the kind of complex interventions integrating multiple strategies that are increasingly the norm. The work by Boyd and colleagues identified 39 unique strategies for each site in their study (6 sites total), while Bunger and colleagues identified 45 unique strategies in their implementation activities (11, 13). In addition, implementation is frequently a multi-phased process, requiring preparatory work, implementation launch, as well as post-implementation activities aimed at increasing reach, adoption, or sustainment (15). And yet most implementation evaluations focus on a single phase of the process, most commonly implementation. This allows for focused examination of core activities and lessons learned, as in a recent study of factors associated with uptake of an evidence-based exercise group for seniors (16), but may constrain the information available on how strategies were used over the full course of implementation (17). This has limited the amount of empirical data available on how the timing of specific strategies, or the sequence in which they are rolled out, may impact the success of implementation. In one novel study attempting to tackle this problem, Yakovchenko and colleagues conducted qualitative comparative analysis of strategies, and identified specific strategy combinations linked to high levels of treatment initiation (18). The authors were unable, however, to discern whether these findings were impacted by the timing or sequence of strategies (18). Similarly, although a handful of studies have examined implementation across multiple phases (11, 13, 15, 19), few have provided significant detail regarding when and how implementation strategies were deployed (20).

The question of how best to document and describe implementation strategies in multi-phase work, therefore, remains salient. In the implementation research described here, we draw upon the Replicating Effective Programs (REP) framework (21) (Figure 1), which functions as an evidence-based roadmap for the implementation of interventions by outlining implementation strategies to be employed across four phases: pre-conditions, pre-implementation, implementation, and maintenance and evolution (21, 22). During the pre-conditions and pre-implementation phases, careful attention is paid to intervention packaging. In the implementation phase, attention is paid to training, technical assistance, and fidelity. And in the maintenance and evolution phase, emphasis is placed on planning and recustomizing for long-term sustainment and spread (22, 23). The “Enhancing Mental and Physical Health of Women Veterans through Engagement and Retention” (EMPOWER) Quality Enhancement Research Initiative (QUERI), funded by the U.S. Department of Veterans Affairs (VA), has undertaken a program of three studies making shared use of REP as an organizing framework (24).

Although the call for greater specificity in describing implementation strategies is important in advancing implementation science, we have found little guidance on how to apply Proctor and colleagues’ recommendations in the context of complex, multi-component interventions, on at least three fronts. First, there is the question of how to ensure all strategies are effectively identified for reporting, given that frameworks such as REP have not previously been described in a manner consistent with newer taxonomies and specification guidelines. Second, use of packaged frameworks such as REP raises questions regarding how to track strategies that may occur at multiple time points, occur in a particular sequence, and/or overlap with other strategies. Similarly, guidance is rarely provided regarding whether component strategies are essential or optional, or their suggested dose or intensity, making it difficult to assess the fidelity with which the framework was followed in resulting trials. Third, evaluating the impact of specific strategies can be difficult, given that implementation outcomes are likely to reflect the cumulative impact of strategies over time.

---

**Figure 1** | Replicating Effective Programs Implementation Strategy*. Enhanced with stakeholder engagement and complexity science. *Adapted from Ref. (21, 22).
In addition, there is a practical challenge associated with operationalizing complex implementation efforts across multiple sites, in ensuring all activities necessary for both implementation and evaluation are occurring at the appropriate time and place. Development of a formal implementation blueprint has been identified as an implementation strategy unto itself, with the suggestion that a blueprint should include the implementation effort’s aim or purpose, intended scope, timeframe, milestones, and appropriate progress measures, and that it should be used and updated over time (9). But while excellent guidelines exist for intervention mapping in health promotion more generally (25), preparing an implementation research proposal (26) or manuscript (27), as well as describing the suggested components of an implementation plan (28), relatively little literature has described how to develop a practicable blueprint for use in organizing the many-tentacled process of implementation evaluation.

To address these concerns, we embarked on a prospective, formative, and iterative team-based process for mapping a multi-component implementation strategy, REP, to recommend taxonomies of implementation strategies. Our primary goal in doing so was to support more effective evaluation of overlapping and sequenced implementation strategies. We also sought to support the operationalization of a complex intervention, providing an implementation blueprint to outline activities and tasks at each phase, and the actors or point persons responsible for those activities. In the current paper, we describe this process alongside the method by which we used the resulting strategy matrix to support development of a formal evaluation plan for one of the EMPOWER studies, “Facilitating Cardiovascular Risk Screening and Risk Reduction in Women Veterans” (known as CV Toolkit), aimed at using a gender-tailored toolkit to reduce cardiovascular (CV) risk among women Veterans in VA primary care settings (24).

METHODS

Implementation Study

The CV Toolkit is comprised of evidence-informed practices aimed at reducing CV risk among patients in primary care and tailored to meet the needs of women Veterans in the VA (Table 1). The CV Toolkit evolved in response to a need for consistent screening and documentation, increased CV risk reduction services and support for women Veterans in VA primary care. REP pre-conditions work leading up to the formal CV Toolkit study included obtaining input from national operations partners and clinical stakeholders regarding potential gaps in women Veterans’ CV risk assessment and care services (24). Pre-conditions work also included focus groups conducted by the study leads (BBM and MF) with primary care providers and women Veteran patients, who identified a variety of barriers and facilitators to effective CV risk management (24). The CV Toolkit was developed as a set of evidence-informed practices intended to address the needs identified by stakeholders and is centered around three specific items: patient education and self-screening of CV risks, provider documentation of CV risks in the electronic health record, and a facilitated group to help patients identify and set behavioral health goals [e.g., the Gateway to Healthy Living program (hereafter, Gateway)]. Gateway is a VA program first piloted in 2015 and now being implemented across VA nationwide, which focuses on motivating and supporting Veterans with chronic conditions such as CV disease or risk conditions to engage in services aimed at reducing their risk (29). Previous evaluation of patient experiences with Gateway suggest high rates of goal setting and linking patients to existing programs, as well as high satisfaction with the Gateway sessions (29). In addition, surveys of staff suggest that the Gateway program was perceived as “very helpful” in connecting Veterans to programs and resources (29).

The CV Toolkit provides a process for assessing women’s CV risk via a patient self-report risk screener, facilitates patient-provider communication and documentation of risk data via a provider-facing computer template embedded in the electronic medical record, and educates providers in shared decision-making and effective clinical action around risk reduction. Women are given the option of participating in women-only Gateway groups, which are tailored for women and focus on CV risk, offer patient education and activation, and serve as an entry point for patients to receive information, goal setting, and referral to other programs and services as needed [additional detail on this and other EMPOWER projects is available (24)].

Having been developed specifically to meet the needs of women Veterans in VA primary care, the CV Toolkit is currently being implemented at two VA facilities with moderately large comprehensive Women’s Health (WH) clinics, with two additional facilities slated for future implementation. Clinics are eligible if they have multiple primary care providers serving women patients (ideally 6 or more providers) and each provider has at least 100 unique women Veteran patients and at least 10%

<p>| TABLE 1 | Summary of Cardiovascular (CV) Toolkit components. |</p>
<table>
<thead>
<tr>
<th>Component</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient education and activation</td>
<td>Educate patients regarding CV risks</td>
</tr>
<tr>
<td>Information sheets, posters, brochures</td>
<td>Inform eligible patients can choose to participate in research component of project (surveys/interviews)</td>
</tr>
<tr>
<td>Opt-in/ Opt-out step</td>
<td>Identify patients with any CV risks; Facilitate patient-provider communication regarding risk factors and appropriate health goals</td>
</tr>
<tr>
<td>CV risk screener</td>
<td>Identify and document patients with CV risks using template and using risk calculator, embedded within the electronic health record; Track use of template by providers on unique patients; Track patient-provider action step or goal</td>
</tr>
<tr>
<td>CV risk computerized template</td>
<td>Educate patients about options for healthy living; provide support in setting behavioral health goals and referrals to appropriate services; Identify and track CV goal set by patient in group</td>
</tr>
<tr>
<td>Gateway to healthy living facilitated group</td>
<td>Follow-up phone calls after Gateway attendance to assess progress with behavioral health goals; Identify barriers to CV goal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Component</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient education and activation</td>
<td>Educate patients regarding CV risks</td>
</tr>
<tr>
<td>Information sheets, posters, brochures</td>
<td>Inform eligible patients can choose to participate in research component of project (surveys/interviews)</td>
</tr>
<tr>
<td>Opt-in/ Opt-out step</td>
<td>Identify patients with any CV risks; Facilitate patient-provider communication regarding risk factors and appropriate health goals</td>
</tr>
<tr>
<td>CV risk screener</td>
<td>Identify and document patients with CV risks using template and using risk calculator, embedded within the electronic health record; Track use of template by providers on unique patients; Track patient-provider action step or goal</td>
</tr>
<tr>
<td>CV risk computerized template</td>
<td>Educate patients about options for healthy living; provide support in setting behavioral health goals and referrals to appropriate services; Identify and track CV goal set by patient in group</td>
</tr>
<tr>
<td>Gateway to healthy living facilitated group</td>
<td>Follow-up phone calls after Gateway attendance to assess progress with behavioral health goals; Identify barriers to CV goal</td>
</tr>
</tbody>
</table>
of their total patient panel is female. Implementation of the CV Toolkit is being evaluated using a non-randomized stepped wedge design to detect differences before and after implementation at each site; this design will also allow for comparisons across sites and providers as the toolkit is implemented (30). The objective of the current work was to develop a step-by-step blueprint operationalizing use of implementation strategies across the CV Toolkit rollout, with the primary goal of guiding evaluation.

Overview/Setting
To develop a comprehensive map of fully specified implementation strategies included as part of the CV Toolkit project, and to link these strategies to our longitudinal evaluation plan, we followed a five-step process, as outlined below. Participants in the strategy mapping process included six team members with overlapping roles central to implementation (including a clinician-researcher who serves as a liaison with sites and provides education for clinicians), intervention (including a health promotion specialist charged with leading Gateway groups and serving as an external facilitator for sites), and evaluation (including experts in health services and implementation research, anthropology, sociology, and biostatistics).

The five-step process includes the following:

(1) Study activity list generation. We first developed a list of CV Toolkit activities as described in the approved human subjects’ protocol, using a previously defined method for treating study documents as primary texts for analysis (31). The CV Toolkit protocol, including the implementation plan, was developed in response to findings from pre-conditions work, and therefore built upon deep knowledge of the VA primary care context and the needs and gaps in care for both women Veterans and their primary care providers. From the beginning, project activities were planned in accordance with the enhanced REP framework, with specific tasks occurring in sequence over the pre-conditions, pre-implementation, implementation, and maintenance and evolution phases. The enhanced REP framework used by EMPOWER QUERI projects (24), building on the original REP framework (21), places more focus on participatory action within complex adaptive systems in VA clinical settings. Once initial activity lists had been generated by two team members (Alexis K. Huynh and Erin P. Finley), these lists were compared and areas of initial discrepancy were discussed with the CV Toolkit Co-Principal Investigators (Bevanne Bean-Mayberry and Melissa M. Farmer) to achieve consensus. The team then categorized each activity as occurring in support of (1) research goals, (2) intervention delivery, or (3) implementation.

(2) Mapping study activities to implementation strategies. Once we had identified all implementation-related activities defined in the protocol, we then mapped these where possible to corresponding implementation strategies, as defined in the Powell compilation (9). As in Step (1), mapping was conducted separately by two team members and then compared, with any discrepancies discussed to consensus with study Co-PIs and other members of the project team, including those providing clinical care in targeted sites and working within the Gateway program. In most cases, the match was clear. Nonetheless, some REP activities did not map to any of the compiled strategies (e.g., collecting data on the timing of implementation launch, which we determined to be a research activity rather than implementation activity), and were not included in the strategy matrix.

(3) Specifying implementation strategies by REP phase and conceptual category. Early in the mapping process, it became clear that certain implementation strategies—e.g., coalition building—were occurring at multiple timepoints over the course of the CV Toolkit study. We therefore took care to specify how and when each strategy would be operationalized during each of the relevant REP phases (see Table 2 for final version) (21). In addition, following Proctor’s recommendations for reporting on use of implementation strategies, we provided full description across each of the seven domains for each strategy, including the actors involved, actions undertaken, targets, dose of the implementation strategy, and anticipated outcomes (3). We also organized the strategies into broader conceptual categories, as proposed by Waltz et al. (10), to evaluate whether specific categories of effort (e.g., stakeholder engagement) emerged at different phases over the course of the study.

(4) Iterative refining of implementation strategy mapping. An initial matrix summarizing the above work was reviewed during a series of team meetings with CV Toolkit study Co-PIs and the overall EMPOWER QUERI PI (Alison B. Hamilton), who provided feedback clarifying the nature, sequence, and/or intent of implementation-related activities. The matrix and mapping process were also presented to larger combined groups of implementation agents and researchers, who offered helpful input regarding how to make the matrix as comprehensive and streamlined as possible. The strategy matrix was iteratively refined from these meetings, resulting in a final matrix (see Table 2) providing detailed description of each implementation strategy planned as part of CV Toolkit implementation. The final strategy matrix was reviewed and validated by the full project team, including members responsible for implementation of the CV Toolkit as well as those tasked with evaluation.

(5) Developing an implementation blueprint. The completed strategy matrix provided a clear step-by-step plan for rolling out implementation strategies to facilitate implementation of the Toolkit, complete with their timing, target, and outcomes. This allowed us to plan for appropriate evaluation of our enhanced REP strategy at each site and across sites. In evaluating the effectiveness of CV Toolkit implementation across this study, we aim to quantitatively assess adoption of three components of the intervention: (1) completion of the CV risk template in the electronic health record by the provider or member of the care team; (2) number of patients who attend the Gateway to Healthy Living facilitated groups and number of follow-up calls made to patients following Gateway attendance; and (3) patient referrals for services. Data captured by the CV risk computer template and other administrative data will be used to examine these outcomes for each provider.
<table>
<thead>
<tr>
<th>REP phase</th>
<th>Pre-condition</th>
<th>Pre-implementation</th>
<th>Implementation</th>
<th>Maintenance and evolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>1. Conduct local needs assessment</td>
<td>1. Establish need for the intervention</td>
<td>1. Discuss CV Toolkit with key stakeholders during site visits</td>
<td>1. Regular communication with opinion leaders throughout intervention (to learn from them what is working and inform them what is not working at other sites).</td>
<td></td>
</tr>
<tr>
<td>2. Inform local opinion leaders</td>
<td>2. Determine feasibility at local site</td>
<td>2. Explain core elements and options for adapting delivery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Develop educational materials</td>
<td>3. Inform local opinion leaders</td>
<td>3. Further local tailoring of educational materials for each site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Promote adaptability</td>
<td>4. Develop educational materials</td>
<td>4. Discuss educational needs of teams at sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Build a coalition</td>
<td>5. Promote adaptability</td>
<td>5. Explain core elements and options for adapting delivery to key stakeholders during site visits</td>
<td>1. Interactive component</td>
<td>1. Collaborate with local teams to develop CV Toolkit Implementation Playbook</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Strategy</th>
<th>Actions by REP phase and month</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>REP phase</strong></td>
<td><strong>Pre-condition</strong></td>
</tr>
<tr>
<td><strong>Month</strong></td>
<td>1</td>
</tr>
<tr>
<td>8. Provide local technical assistance</td>
<td>1. Discuss educational needs of teams at sites</td>
</tr>
<tr>
<td>10. Identify and prepare champions</td>
<td></td>
</tr>
<tr>
<td>11. Assess for readiness, &amp; and identify barriers &amp; and facilitators</td>
<td></td>
</tr>
<tr>
<td>12. Develop formal implementation blueprint</td>
<td></td>
</tr>
<tr>
<td>13. Audit and provide feedback</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
### TABLE 2 | Continued

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Actions by REP phase and month</th>
</tr>
</thead>
<tbody>
<tr>
<td>REP phase</td>
<td>Pre-condition</td>
</tr>
<tr>
<td>Month</td>
<td>1</td>
</tr>
<tr>
<td>14. Purposefully reexamine the implementation</td>
<td></td>
</tr>
<tr>
<td>15. Conduct cyclical small tests of change</td>
<td></td>
</tr>
<tr>
<td>16. Develop an implementation glossary</td>
<td></td>
</tr>
</tbody>
</table>

- Develop stakeholder interrelationships
- Train and educate stakeholders
- Use of evaluative and iterative strategies
- Adapt and tailor to context
- Provide interactive assistance
at each site and will allow us to assess whether and how adoption varies as strategies are enacted over the course of implementation. It is anticipated that successful adoption of CV Toolkit will also impact patient–provider communication and patient experiences of and engagement with care. We are therefore collecting qualitative data regarding patients' and providers' experiences of and engagement with CV Toolkit implementation, including adoption, acceptability, feasibility, engagement, and satisfaction (32). We are also conducting reflective discussions with team members to aid in documenting when and how key implementation activities occur (33). Taken in sum, these data will be integrated to allow for process and summative evaluation (see Table 3), as described in the published protocol (24).

<table>
<thead>
<tr>
<th>Replicating Effective Programs phase*</th>
<th>Phase 1: pre-conditions</th>
<th>Phase 2: pre-implementation</th>
<th>Phase 3: implementation</th>
<th>Phase 4: maintenance and evolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>11 12 13 14 15 1 2 3 4</td>
</tr>
<tr>
<td>Provider and administrator interviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Phase 1: intervention planning, needs assessment, and acceptability;</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Phase 2: factors likely to affect adoption, acceptability, feasibility, satisfaction, penetration/reach.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Phase 4: experiences of intervention/implementation; adaptations made in practice; suggestion for future adaptations to inform effectiveness and spread.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provider surveys</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>• Measuring organizational readiness for patient engagement (more)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient interviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Phase 3: factors likely to affect adoption, acceptability, feasibility, satisfaction, penetration/reach.</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Phase 4: experiences of intervention/implementation; challenges, problem-solving, and suggestions for change/adaptation.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Surveys (pre- and post-intervention)</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>• Primary outcomes: program engagement and retention; change in targeted symptom or risk reduction behavior;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Secondary outcomes: satisfaction (at f/u only), global health, out of role days;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Potential moderators: engagement, patient demographics, social support, mental health</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Periodic reflections (discussions with team members to document)</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>• History and trajectory of implementation events</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Activities and interrelationships, including stakeholder engagement;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Adaptations to intervention components and/or implementation strategies;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Contextual factors with potential impact for implementation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrative data</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>• Referral monitoring</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>• Patient engagement</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>• Patient outcomes</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Text analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Review of changes occurring to intervention components and/or implementation strategies per T1 (baseline) proposal materials and subsequent institutional review, amendments, and other study documentation</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

*At each implementation site, phases are expected to occur as follows: pre-conditions (6 months); pre-implementation (6 months); implementation (15 months); maintenance and evolution (4 months).
As a means of verifying expected links between intervention components, implementation strategies, and outcomes of interest, we conducted a process of simulating data. Following the example of Zimmerman and colleagues (34), who suggest use of modeling to aid in implementation planning, we first mapped the flow of patients attending the women’s health primary care clinic and the process by which they receive referrals to the Gateway. Walking through the expected flow of patients in clinic with the study team, we estimated the likelihood of the provider completing the computer template, and making referral to Gateway; estimates were allowed a range of likelihood (e.g., 5–20%) to provide a lower and upper bound. We also estimated a rate of increase in these activities as the implementation period progressed. Estimates were intended to be conservative and were based in the team’s clinical and research experience of VA Women’s Health primary care clinics and change initiatives. Walking through the simulation process prompted useful discussion regarding where barriers and “bottlenecks” were likely to occur, stimulating discussion of how best to work with frontline providers and staff in overcoming those barriers. Final estimates were used to populate and refine a draft of a pragmatic implementation and evaluation blueprint that stipulates the general timing of activities and data collection, aids in assessing implementation outcomes, and ensures effective coordination of implementation and research activities (Figure 2). Strategy mapping activities occurred over the course of a one-year pre-implementation period during which other preparatory activities were ongoing, including identification of sites and site needs assessment and tailoring.

RESULTS

Table 4 below enumerates the 16 discrete implementation strategies intended for use as part of the CV Toolkit’s implementation effort according to enhanced REP. Strategies fell into five main categories, primarily related not only to use of evaluative and iterative strategies (6) and development of stakeholder interrelationships (5), but also reflecting efforts to train and educate stakeholders (2), adapt and tailor to context (2), and provide interactive assistance (1).

Table 3 delineates planned use of strategies across each of the four REP phases. Four of the 16 strategies identified are to be deployed during a single REP phase: conduct local needs assessment in the pre-condition phase; assess for readiness and identify barriers and facilitators in pre-implementation; conduct cyclical small tests of change during implementation; and develop an implementation glossary during maintenance and evolution. All other strategies occurred across more than one phase of the implementation effort.

Most (9 out of 16) strategies are initiated in the pre-condition phase. These nine are varied and include the following: involve executive boards; build a coalition; inform local opinion leaders;
TABLE 4 | Number of implementation strategies by conceptual cluster (10).

<table>
<thead>
<tr>
<th>Strategy conceptual cluster</th>
<th>Frequency</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop stakeholder interrelationships</td>
<td>5</td>
<td>Involve executive boards Build a coalition Inform local opinion leaders Identify and prepare champions Develop an implementation glossary</td>
</tr>
<tr>
<td>Use evaluative and iterative strategies</td>
<td>6</td>
<td>Conduct local needs assessment Conduct cyclical small tests of change Assess for readiness and identify barriers and facilitators (local resources) Develop formal implementation blueprint Audit and provide feedback Purposely reexamine the implementation</td>
</tr>
<tr>
<td>Train and educate stakeholders</td>
<td>2</td>
<td>Conduct educational meetings Develop educational materials</td>
</tr>
<tr>
<td>Adapt and tailor to context</td>
<td>2</td>
<td>Tailor strategies Promote adaptability</td>
</tr>
<tr>
<td>Provide interactive assistance</td>
<td>1</td>
<td>Provide local technical assistance</td>
</tr>
</tbody>
</table>

conduct local needs assessment; develop educational materials; conduct educational meetings; tailor strategies; promote adaptability; and provide local technical assistance. By contrast, there are fewer implementation strategies initiated in the remaining REP phases: two in the REP pre-implementation phase (identify and prepare champions and assess for readiness and identify barriers and facilitators), four in the implementation phase (conduct cyclical small tests of change, develop formal implementation blueprint, audit and provide feedback, and purposely reexamine the implementation), and one in the maintenance and evolution phase (develop an implementation glossary).

Strategies occurring in later REP phases focus on two main categories of activity: use of evaluative and iterative strategies and developing stakeholder interrelationships.

Once initiated, most strategies (12 of the 16) are to be deployed during multiple REP phases. For example, strategies that involve training and education of stakeholders (e.g., developing educational materials and conducting educational meetings) are deployed during pre-condition, pre-implementation, and implementation phases, as are strategies for informing local opinion leaders, providing local technical assistance, and identifying and preparing champions. Most strategies that involve use of evaluative and iterative strategies (e.g., developing formal implementation blueprint, audit and provide feedback, and purposely reexamine the implementation) are to be deployed during implementation and maintenance and evolution phases. Strategies for promoting adaptability are deployed during the latter three REP phases (pre-implementation, implementation, and maintenance and evolution), while strategies for building a coalition occur across pre-conditions, pre-implementation, and maintenance and evaluation phases. Finally, two of the strategies (tailor strategies and involve executive boards) are deployed during all four REP phases.

Results for the implementation scenario simulations are presented in Figure 2. Figure 2A includes outcomes related to providers’ entry of CV risk screener data into the medical record and referrals to VA programs. Figure 2B models attendance at Gateway groups and follow-up phone calls to Gateway participants. Team members hypothesized that providers would enter patient screener information into the CV template during patient appointments 15% of the time during early implementation. Team members expected improvements in the proportions of providers entering the information over time, such that at the end of 18 months of implementation, the proportion would increase to 35%. Second, team members hypothesized that referrals by providers to other VA services would increase by 15% by the end of implementation. Based on these parameters, approximately up to 21% of patients were expected to be receiving any new referrals by the end of the study period. Team members hypothesized that Gateway participation would increase to 30% and most participants would receive follow-up phone calls by the end of implementation.

DISCUSSION

Recent guidelines for specifying implementation strategies raise challenges for implementation efforts making use of multiple or packaged strategies, such as the use of enhanced REP in the EMPOWER QUERI. These challenges include how best to describe each individual strategy and its components, develop a practical blueprint for operationalizing implementation and research activities, and ultimately, plan for a program evaluation that takes the cumulative impact of packaged strategies into account. We conducted a prospective, formative, and iterative process of strategy mapping to address these challenges, mapping implementation activities and strategies into an explicit blueprint by implementation phase and conducting a simulation exercise with project team members to validate our evaluation plan. The blueprint articulates the projections of what we anticipate in implementing the CV Toolkit, and serves as an accounting tool that allows us to track and compare our projections to on-the-ground implementation progress as we carry out the intervention. The method of mapping has provided new insight into where, when, and how each strategy is deployed, allowing us to formulate a targeted multi-method evaluation plan.

We identified five categories of strategies to be used in the implementation of the CV Toolkit: use of evaluative and iterative strategies, develop stakeholder interrelationships, adapt and tailor to context, train and educate stakeholders, and provide interactive assistance. These five categories correspond to the five that Waltz and colleagues rated as having the highest importance in achieving successful implementation (10). Communication, an additional category of strategies suggested by Boyd and colleagues (13), appeared to emerge in these data as an essential component of nearly all strategies, rather than a distinct category unto itself. We also mapped evaluative and iterative strategies as occurring most frequently in the CV Toolkit implementation, an emphasis that appears to be supported by Waltz and colleagues’ rating of evaluative and iterative
strategies as the single most important category of strategies. It is noteworthy that explicitly financial strategies are not used in the CV Toolkit. This contrasts with the work of Honeycutt and colleagues, who identified financial and technical assistance as effective mechanisms for dissemination of evidence-based programs (35). Similarly, Cunningham and Card found that funding, staff, and other resources was the only factor significantly associated with implementation of evidence-based interventions (17). In future work, it will be important to compare how financial strategies affect implementation in integrated versus decentralized healthcare systems (36).

In addition to identifying the relative frequency of strategies, mapping the list of discrete strategies to be used across REP phases provided significant insight into the timing of when strategies are used in this project, and to what ends. For example, although evaluative and iterative strategies are the most frequently occurring, these strategies occur primarily during implementation and maintenance and evolution phases. By contrast, most other strategies are initiated in the pre-conditions phase, thus underscoring the importance of the early phase in laying the groundwork for large-scale implementation studies. Our current study is similar to other implementation evaluation studies that examine implementation by phases, such as that by Chamberlain and colleagues, who focused on two implementation strategies and found that sites ceased progress during pre-implementation phase (15). Similarly, Blackford and colleagues (19) have also made use of an evaluation tool to track progress in implementing an advance care planning initiative, finding the tool useful in supporting planning, tracking progress, and providing direction for future change. In all, our current study and those in the literature speak to the importance of timing in evaluating how differing strategies support effective implementation.

We found dose to be the most difficult domain to define for 12 of the 16 strategies mapped, and specifically for those strategies deployed across multiple REP phases. Issues to be resolved include how to quantify dose for each strategy (e.g., unit of analysis), the relationship between length of time and intensity of effort involved in calculating dose, and what activities “count” as deployment of a strategy, e.g., if a strategy is used only briefly or mentioned in an email. Additional issues that arose include how best to quantify the cumulative effects of strategies deployed at multiple phases, e.g., additively or multiplicatively. These issues hold true for all strategies except for the four that we identified as being deployed during a single REP phase, which are more easily counted and tracked as activities. In pragmatic implementation, it may not always be feasible or practical to specify every component of implementation strategies when working with complex, multi-component packages. The literature points to differing approaches as to how to define dosage in implementation evaluation studies. For example, Boyd and colleagues operationalized dose as intent to use strategies (13). Similarly, Ferm and colleagues defined dose in terms of intervention fidelity (i.e., number of sessions of the intervention compared to the number of sessions that was supposed to be delivered). By contrast, Bunker and colleagues (11) operationalized dose in terms of person-hours invested in implementation. Honeycutt and colleagues (35) found that sites implementing had different interpretations of defining completion of core elements and suggested that future studies might benefit from explicit guidance on quantifying dose of program core elements. Nonetheless, the recent guidelines by Powell, Proctor and colleagues encourage thoughtful attention to these components.

Simulating the implementation scenarios in which the CV Toolkit is deployed was helpful because it served as a “run-through” of our evaluation plan. We identified the many moving and interacting components of the Toolkit and how each is likely to contribute to the outcomes of interest. We also clarified the information that we can expect to collect routinely over time and across sites, which we expect to serve as parameters and data for longitudinal analyses. The simulation exercise also served to validate our evaluation plan that explicitly accounts for the multi-level structure of the data, taking into consideration the context-dependent nature of implementing the Toolkit.

We believe there are a number of advantages to the strategy mapping approach described here. This method provides a low-burden process for achieving specification of strategies. It also supports developing an implementation blueprint and comprehensive evaluation plan, with potential for examining adherence. We also believe that strategy mapping is likely to be easier and more supportive of effective implementation if done prospectively rather than retrospectively. Mapping is likely to be fruitful in ensuring that all elements of an implementation research effort—including the intervention, implementation plan, and evaluation plan—have been clearly articulated prior to launch. In the case of the CV Toolkit project, the mapping process has provided structure for implementation by allowing for detailed front-end specification of project activities, development of a succinct but comprehensive blueprint for activities across each of the four REP phases, and simulation of the longitudinal quantitative data likely to emerge across sites, thus providing both guidelines for and an opportunity to “test-run” implementation and evaluation activities. Visual representation of planned strategy rollout can also serve as a tracking tool to support identifying where the project, or a specific site, deviates from the expected use of or sequencing of strategies. Mapping strategies helps to organize, plan, and clarify the implementation process by specifying the necessary action steps per phase, and milestones along the implementation timeline. Moreover, mapping implementation strategies allows us to identify and prioritize key strategies that we can leverage to improve outcomes. Finally, as we move forward with CV Toolkit implementation, in partnership with local and national stakeholders, we expect that strategy mapping will also support development of implementation playbooks (37)—i.e., brief primers providing “how to” or “lessons learned” information—intended to facilitate more rapid dissemination, scale-up, and spread.

Potential disadvantages of this approach include the fact that it requires substantial time during the initial project planning phases. We conducted the activities described over a one-year period preparatory to implementation launch; however, we believe this process could be conducted much more rapidly following the outline offered here. Although mapping strategies
The CV Toolkit examined directly in the simulation exercise described above, we are also using semi-structured interviews to assess acceptability, feasibility, and satisfaction among patients receiving the CV Toolkit and providers and staff members delivering it. However, we believe that strategy mapping preparatory to implementation is likely to make tracking easier and potentially more accurate by functioning as a practical checklist for expected activities that allows for the benchmarking of implementation progress.

Future research should continue to explore the utility of this and other methods for mapping strategies in complex implementation. One interesting possibility for this work is likely to involve a more participatory approach, working directly with sites and other stakeholders to delineate key strategies and plan for pragmatic evaluation. The role of data capture in providing information on whether and when adoption is occurring provides the opportunity to further explore how best to observe, track, and communicate with stakeholders regarding implementation progress and outcomes (38). We are continuing to explore questions related to the analytic utility of strategy mapping as we proceed with the multi-site CV Toolkit study, including whether the process can be used to identify core components of packaged strategies like our enhanced REP, whether specific categories of strategies appear to be associated with specific outcomes [similar to the approach used by Boyd et al. (13)], and whether differing combinations or sequences of strategies appear to be associated with differential outcomes [similar to the findings by Yakovchenko (18)]. Notably, as illustrated in Table 3, our evaluation plan is multi-method and integrates both quantitative and qualitative data sources to address these research questions. For example, in addition to the questions related to adoption and reach of the CV Toolkit examined directly in the simulation exercise described above, we are also using semi-structured interviews to assess acceptability, feasibility, and satisfaction among patients receiving the CV Toolkit and providers and staff members delivering the CV Toolkit in their clinics.

CONCLUSION

We update recent guidance on specification of implementation strategies by considering the implications of such guidance for use of multi-strategy frameworks such as enhanced REP, and propose a novel method to support strategy mapping in complex interventions, with the goal of facilitating both implementation and evaluation efforts. Our strategy mapping approach is innovative in offering a clear and structured method for stipulating when and how implementation strategies occur across the entire life cycle of an implementation effort, in this case across the four REP phases. By doing so, the method aids in fully documenting how implementation activities proceed, to support more effective description and replicability where implementation proves successful. This method also aids in developing plans for evaluation and analysis by clarifying the timing of events and where specific implementation strategies are occurring singly or in combination. Our results identified interesting patterns in the sequence of strategies, particularly related to the importance of pre-implementation activities in laying the groundwork for implementation, as well as the differing ways that specific implementations strategies may be used across different REP phases (e.g., with coalition partners providing support for local uptake during early phases and informing strategies for dissemination and spread in later phases). This approach may therefore be of particular usefulness in implementation efforts employing multi-phase frameworks, such as EPIS (23). Ultimately, understanding timing of implementation strategies will aid in the summative evaluation that utilizes the non-randomized stepped wedge design that explicitly accommodates for the naturalistic roll-out of interventions and programs. Furthermore, specifying strategies into their functional components provides a level of detail on implementation activities that is likely to aid in identifying not only whether the overall implementation has been successful in impacting clinical and patient outcomes, but also by what mechanisms. Finally, in operationalizing and specifying the implementation strategies used in each phase of implementation, we seek to advance understanding of how implementation strategies—individually and in combination—function to support effective practice change. The work presented here provides a model for developing comprehensive implementation and evaluation blueprints to support the increasing methodological complexity of work being done in implementation science.

AUTHOR CONTRIBUTIONS

AKH developed the method, analyzed, synthesized, and interpreted the findings, and drafted and critically revised the manuscript. ABH conceived the design of the overall project and manuscript, provided feedback on the method, interpretations of implementation, and research activities, interpreted the findings, and drafted and critically revised the manuscript. BB-M provided feedback on the method, interpretations of implementation, and research activities, interpreted the findings, and drafted and critically revised the manuscript. MF provided feedback on the method, interpretations of implementation, and research activities, interpreted the findings, and drafted and critically revised the manuscript. SS provided feedback and interpretations of implementation and research activities, interpreted the findings, and drafted and critically revised the manuscript. EF developed the method, analyzed, synthesized, and interpreted the findings, and drafted and critically revised the manuscript.

ACKNOWLEDGMENTS

The views expressed in this manuscript are those of the authors and do not reflect the position or policy of the Department of Veterans Affairs or the United States Government. Versions of this paper were presented in 2017 at the 4th Biennial Society for Implementation Research Collaboration (SIRC) in Seattle,
REFERENCES


FUNDING

The EMPOWER implementation initiative described was funded through VAs Quality Enhancement Research Initiative (QUERI) (grant number 15–272), which uses operational funds to support program improvement. The CV Toolkit project is considered research and was approved by the Central VA Institutional Review Board and local site Research and Development Boards.


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Implementation Mapping: Using Intervention Mapping to Develop Implementation Strategies

Maria E. Fernandez*, Gill A. ten Hoor3, Sanne van Lieshout3, Serena A. Rodriguez1,4, Rinad S. Beidas5,6, Guy Parcel1, Robert A. C. Ruiter2, Christine M. Markham1 and Gerjo Kok2

1 Center for Health Promotion and Prevention Research, University of Texas Health Science Center at Houston School of Public Health, Houston, TX, United States, 2 Department of Work and Social Psychology, Maastricht University, Maastricht, Netherlands, 3 Department of Public Health, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands, 4 Department of Population and Data Sciences, University of Texas Southwestern Medical Center, Dallas, TX, United States, 5 Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, United States, 6 Department of Medical Ethics and Health Policy, University of Pennsylvania, Philadelphia, PA, United States

Background: The ultimate impact of a health innovation depends not only on its effectiveness but also on its reach in the population and the extent to which it is implemented with high levels of completeness and fidelity. Implementation science has emerged as the potential solution to the failure to translate evidence from research into effective practice and policy evident in many fields. Implementation scientists have developed many frameworks, theories and models, which describe implementation determinants, processes, or outcomes; yet, there is little guidance about how these can inform the development or selection of implementation strategies (methods or techniques used to improve adoption, implementation, sustainment, and scale-up of interventions) (1, 2). To move the implementation science field forward and to provide a practical tool to apply the knowledge in this field, we describe a systematic process for planning or selecting implementation strategies: Implementation Mapping.

Methods: Implementation Mapping is based on Intervention Mapping (a six-step protocol that guides the design of multi-level health promotion interventions and implementation strategies) and expands on Intervention Mapping step 5. It includes insights from both the implementation science field and Intervention Mapping. Implementation Mapping involves five tasks: (1) conduct an implementation needs assessment and identify program adopters and implementers; (2) state adoption and implementation outcomes and performance objectives, identify determinants, and create matrices of change objectives; (3) choose theoretical methods (mechanisms of change) and select or design implementation strategies; (4) produce implementation protocols and materials; and (5) evaluate implementation outcomes. The tasks are iterative with the planner circling back to previous steps throughout this process to ensure all adopters and implementers, outcomes, determinants, and objectives are addressed.

Discussion: Implementation Mapping provides a systematic process for developing strategies to improve the adoption, implementation, and maintenance of evidence-based interventions in real-world settings.

Keywords: implementation, dissemination, adoption, intervention mapping, adaptation, implementation strategies, mechanisms of change, health promotion
INTRODUCTION

The ultimate impact of health innovations depends not only on the effectiveness of the intervention, but also on its reach in the population and the extent to which it is implemented properly. The research to practice translation process includes the development of interventions, testing their effectiveness, and ensuring they are adopted, implemented, and maintained over time. However, many research findings are never translated into policy and/or practice, or are done so very slowly, often years after its evidence has been established and with variable levels of implementation and maintenance (3). Program users do not always implement a program as it was intended, leaving out certain elements, or making alterations without careful consideration. This can compromise completeness and fidelity of implementation and subsequently program effectiveness (4, 5). Failing to appropriately implement effective interventions, guidelines, or policies severely limits the potential for patients and communities to benefit from advances in health promotion, medicine, and public health.

In the last decade, implementation science has emerged as the potential solution to this major problem (6, 7). Implementation science refers to the scientific study of methods to increase the adoption, implementation, and maintenance of evidence-based practices, programs, policies, and guidelines (6, 7). The implementation science field provides various implementation theories, frameworks, and models (8, 9). These aim to describe the process of translating research into practice, understand, or explain determinants of implementation, or to evaluate implementation (8).

Despite the rapidly increasing wealth of implementation science insights and knowledge, the majority of programs still fail to systematically plan for adoption and implementation. Instead of planning for all implementation steps from the beginning (i.e., adoption, implementation, and maintenance), the identification or development of implementation strategies typically occurs after the evidence-based intervention has already been developed or following failed implementation efforts (3, 10, 11). There seems to be a high standard for developing interventions to impact health outcomes, but less rigor and thoughtfulness in developing the implementation strategies needed to deliver the intervention. Implementation strategies are methods or techniques used to improve adoption, implementation, sustainment, and scale-up of interventions (1, 2, 12). These strategies vary in their complexity, from discrete or single component strategies to multi-component or bundled approaches (2, 7). They include both the small-scale strategies to influence specific determinants and of a implementation task, and overall packages of strategies influencing adoption, implementation, and maintenance behaviors that will ultimately determine whether a program is adopted, used, and maintained over time (3–8, 11, 13). Problems related to the development and selection of implementation strategies are evident in the literature and include: little use of theory in planning or selecting implementation strategies, lack of explicit articulation of implementation goals, limited understanding of the determinants of implementation to inform strategy development and scant descriptions of the underlying mechanisms of change that are hypothesized to cause the desired effect (14–16). For example, in a study by Davies et al. that reviewed 235 studies, authors reported that only 23% used theory to inform design of implementation strategies (14).

Nevertheless, the field has made significant strides in understanding and categorizing implementation strategies described in the published literature (7) and has suggested general approaches for selecting and describing strategies used (1, 17). These efforts have greatly advanced the field of implementation science. Still, there is little guidance on how to systematically select or plan implementation strategies at multiple ecologic levels to increase adoption, implementation, and sustainability of evidence based interventions nor how to effectively use implementation science theories and frameworks to inform the process. Thus, although useful for better understanding the types of implementation strategies that have been used, the existing inventories do little for program planners attempting to identify the most effective implementation strategies given a complex set of conditions and determinants influencing program use (1).

Researchers and practitioners alike are often forced to plan, develop, or select implementation strategies with very little information about what might work and little consideration about the mechanisms underlying potential change (18, 19). To move the implementation science field forward and to close the research-to-practice gap, a systematic process is needed to help plan for dissemination and implementation of evidence-based interventions that considers determinants, mechanisms, and strategies for effecting change. In this paper, we describe how Intervention Mapping is used to plan or select implementation strategies, a process we call Implementation Mapping.

INTERVENTION MAPPING

Intervention Mapping is a protocol that guides the design of multi-level health promotion interventions and implementation strategies (13). Since its inception, a key feature of Intervention Mapping (Step 5) has been its utility for developing strategies to enhance the adoption, implementation, and maintenance of clinical guidelines (13) and evidence-based interventions (20–26).

Intervention Mapping consists of six steps: (1) conduct a needs assessment or problem analysis by identifying what, if anything, needs to be changed and for whom; (2) create matrices of change objectives by crossing performance objectives (sub-behaviors) with determinants; (3) select theory-based intervention methods that match the determinants, and translate these into strategies, or applications, that satisfy the parameters for effectiveness of the selected methods; (4) integrate the strategies into an organized program; (5) plan for adoption, implementation, and sustainability of the program in real-life contexts by identifying program users and supporters and determining what their needs are and how these should be fulfilled; (6) generate an evaluation plan to conduct effect and process evaluations to measure program effectiveness (13).
Essentially, Steps 1–4 focus on the development of multilevel interventions to improve health behaviors and environmental conditions, Step 5 focuses on the development of implementation strategies to enhance program use, and Step 6 is used to plan the evaluation of both the program itself and its implementation.

Intervention Mapping can advance the field of implementation science via three distinct, yet interrelated, ways. First, the use of Intervention Mapping helps “design for dissemination” (27) a concept that means considering implementation during the development of the intervention. Intervention Mapping does so by guiding planners through a systematic process that engages stakeholders in the development of a program, policy, or practice that is likely to be both effective and usable. Second, IM can be used to systematically adapt existing evidence-based interventions to align them with new populations, geographic regions, or implementation contexts. Third, and most relevant for this paper, Intervention Mapping can help planners to develop, select, or tailor implementation strategies to increase adoption, implementation, and sustainability. Since its inception, a key feature of IM, has been its utility for developing implementation strategies to enhance the adoption, implementation, and sustainability (20–26), nevertheless, its utility has only recently been recognized by implementation scientists (12, 15, 17, 27). Thus, using Intervention Mapping for initial program development, for program adaptation, and/or for planning implementation can reduce the gap between the development of effective clinical practices and programs and their actual use in healthcare settings and communities (28).

Depending on what the evidence-based intervention is that will be implemented, a planner may choose to use all six steps of Intervention Mapping starting with Step 1, or simply Step 5. The distinction lies in whether or not there is an existing “intervention.” If, for example, the task is to develop an intervention to implement clinical practice guidelines at multiple levels of an organization (e.g., changing patient and provider behavior) and/or there are no specific products (activities, training, materials) to be implemented yet, planners should start with Step 1 of Intervention Mapping because they are developing a multi-level intervention that will, in turn, need to be implemented. If, however, there is an existing evidence-based intervention (at one or more levels) that has been developed and tested, planners can focus on how to get this intervention adopted, implemented, and maintained and begin with Intervention Mapping Step 5. Intervention Mapping Step 5 is what we refer to as Implementation Mapping.

**IMPLEMENTATION SCIENCE + INTERVENTION MAPPING = IMPLEMENTATION MAPPING**

Implementation Mapping includes insights from both the implementation science field and from Intervention Mapping. In Implementation Mapping described here, we expand on the four tasks associated with Intervention Mapping Step 5 (identify program implementers, state outcomes and performance objectives for program use, construct matrices of change objectives, design implementation strategies) (13). Here we provide additional details for selecting and developing implementation strategies. Implementation Mapping involves five specific tasks: (1) conduct a needs assessment and identify program adopters and implementers; (2) state adoption and implementation outcomes and performance objectives, identify determinants, and create matrices of change objectives; (3) choose theoretical methods and select or design implementation strategies; (4) produce implementation protocols and materials; and (5) evaluate implementation outcomes. The five tasks are iterative with the planner circling back to previous tasks throughout to ensure all adopters and implementers, outcomes, determinants, and objectives are addressed; see Figure 1.

**Task 1. Conduct an Implementation Needs Assessment**

In Implementation Mapping Task 1, planners conduct (or describe results of) a needs and assets assessment. This is sometimes referred to as identification of barriers and facilitators of implementation. Here we involve all agents including adopters, implementers, and those responsible for maintaining the evidence-based intervention in processes to identify actions needed to implement the program and determinants (barriers and facilitators) of implementation. Ideally this should have happened in Intervention Mapping step 1, but very often, a program planner has insufficient information about the implementation setting and process before the interventions has been developed.

Often, the identification and engagement of implementers occurs late in the intervention development process after the
intervention is developed or proven successful and sometimes even after an implementation strategy has been selected. This can lead to low levels of implementation and maintenance because the selected strategy does not address the most salient determinants of implementation, because it does not fit well within the context, or other reasons. Therefore, the first task in Implementation Mapping is to identify program adopters and implementers. During an intervention development effort adopters and implementers may have already been a part of the team. Planners should ensure that all adopters and implementers important to implementation have been identified and are (still) involved. The questions that need to be answered at the end of task 1 are: (a) Who will decide to adopt and use the program? (b) Which stakeholders will decision makers need to consult? (c) Who will make resources available to implement the program? (d) Who will implement the program? (e) Will the program require different people to implement different components? And (f) Who will ensure that the program continues as long as it is needed (13)?

The identified stakeholders are not only stakeholders at the individual level, but also at all environmental levels. The results of the needs and assets assessment often highlight the need to target multiple adopters and implementers within an implementation setting. For example, while adopters may sometimes also be responsible for program implementation, this is not always the case. Clinic administrators may choose to adopt an evidence-based intervention to improve patient outcomes while physicians, nurses, and other staff are responsible for implementing the intervention with patients. For complex interventions, there may be different adopters and implementers for program components at different levels (clinic or school level vs. provider or teacher level). At the individual level, adopters’ or implementers’ attitudes toward innovations or new programs can influence decisions to adopt or implement the program. Alternatively, at the organizational level, a clinic may lack resources or personnel to implement new systems or protocols. To identify all actors and potential barriers and facilitators to implementation, the needs assessment is essential, and may require initial brainstorms within the implementation planning group and literature reviews, but also interviews with potential adopters and implementers or observations within the setting.

Wallerstein and Duran (29) describe the potential for Community-based Participatory Research (CBPR) to ensure that efforts to understand and improve implementation strategies promote reciprocal learning and incorporate community theories into these efforts. Implementation Mapping emphasizes the application of principles and processes of community based participatory research and planning and engagement of stakeholders at multiple levels. These “core processes,” described as such in the original Intervention Mapping protocol, are fundamental throughout the course of planning implementation strategies and particularly when conducting an assessment of implementation barriers, facilitators, needs, and resources. Including individuals who may adopt, implement, or use the program in understanding contextual and motivational issues and in planning and selecting implementation strategies can help address issues by ensuring integration of the local community’s or clinic’s priorities, perspectives, and practices (29). This also helps ensure that materials, methods, and strategies fit the local context (30–32). Additionally, creating a program in partnership with a community can help leverage community networks for implementation and dissemination (33). Thus, we encourage use of a participatory approach to implementation planning that includes potential adopters, implementers, and maintainers in implementation planning from the beginning of the planning process (34). Consistent with Diffusion of Innovation Theory, Implementation Mapping also encourages the use of a linkage system in which new change agents, program champions, and representatives of those with actual responsibility for implementation are included in the planning group (31, 35, 36). This engagement is essential to gain a realistic understanding of what organizational resources, staffing, financial, and other factors are needed for implementation (29).

Cabassa et al. (37) used Intervention Mapping combined with a community-based participatory planning approach to adapt a healthcare manager intervention focused on improving health of Hispanics with serious mental illness. They used a community advisory board with researchers and stakeholders to review the original intervention and make initial modifications. To ensure that adaptations were acceptable, they then conducted patient focus groups and stakeholder interviews. Following further adaptation based on input, they then used Intervention Mapping to develop an implementation plan, and conducted a pilot study to assess intervention feasibility, acceptability, and preliminary effectiveness. The authors highlight the differences between traditional knowledge translation approaches and a CBPR approach to implementation where stakeholders and partners participate collaboratively to understand and create strategies to improve implementation (38). They used Intervention Mapping to guide the process.

Task 2. Identify Adoption and Implementation Outcomes, Performance Objectives, Determinants, and Change Objectives.

In Implementation Mapping task 2, implementation planners state adoption and implementation outcomes and performance objectives, identify determinants, and develop matrices of change objectives. Outcomes are specific to each adopter and implementer. If adoption and implementation involve multiple actors such as administrators, physicians, and patient navigators, each may have their own adoption and implementation outcomes or performance objectives depending on their role. Performance objectives are essentially the tasks required to adopt, implement, or maintain a program. Adoption and implementation outcomes are often straightforward and simply state the key actor or actors and the adoption, implementation, or maintenance goal. Table 1 lists the adoption and implementation outcomes of the Peace of Mind program, an intervention to increase mammography screening among patients of community health centers. Table 1 also provides examples of outcomes from the Long Live Love
program, a curriculum about love, relationships, and sexuality for secondary schools and vocational schools (see also https://www.langleveldeliefde.nl/docenten/english). After identifying adoption and implementation outcomes, planners state performance objectives for each outcome. Performance objectives, shown in Table 1, are the specific steps, or sub-behaviors, that adopters and implementers must perform to meet the overall adoption and implementation outcomes (13). Performance objectives make clear “who has to do what” for the program to be adopted, implemented, and continued. Performance objectives are action oriented and do not include cognitive processes such as “know” or “believe.” For adopters, the question is: "What do [adopters] have to do in order to make the decision to use [the program]?" These actions may, for example, include comparing the new evidence-based intervention to existing practices, gathering feedback and support from potential implementers, or signing a formal agreement to adopt.

To create performance objectives for implementers, we ask: "What do the program implementers need to do to deliver the essential program components?" Implementation performance objectives may include attending trainings, gathering materials, or updating protocols; Table 1 contains examples of performance objectives from existing projects. And for those responsible for program continuation: "What do they need to do to maintain the program?" Posing these questions may seem obvious, however, they help the planner articulate the exact actions required to put a health promotion intervention into use, details that are not always clear when seeking to develop or select implementation strategies. Answers to these questions are often informed by the needs assessment. Findings from the needs assessment

---

**TABLE 1 | Implementation outcomes and performance objectives: select examples.**

**Program: Peace of mind** ([23, 28])

**Setting: Clinic-based**

<table>
<thead>
<tr>
<th>Target: role</th>
<th>Adoption, Implementation, and Maintenance Outcomes</th>
<th>Performance objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinic decision maker: Adopter</td>
<td>The management team at clinic decides to adopt the Peace of Mind program (PMP) as indicated by the clinic director signing a memorandum of understanding.</td>
<td>1. Agree to participate in PMP 2. Agree to expand mammography services 3. Agree to participate in evaluation 4. Provide a program champion 5. Gain support from stakeholders</td>
</tr>
<tr>
<td>Patient navigator: Implementer</td>
<td>The patient navigator will complete PMP telephone counseling with eligible patients and complete appointment reminder calls.</td>
<td>1. Search schedule for upcoming appointments 2. Conduct telephone barrier counseling 3. Make three attempts to reach patient via phone before appointment</td>
</tr>
<tr>
<td>Program champion: Maintainer</td>
<td>The program champion will ensure clinic leadership maintains PMP as part of the clinic’s standard practice for every appointed mammography patient after initial funding is withdrawn.</td>
<td>1. Discuss with decision makers the continuation of the PMP after funding 2. Work with decision makers to continue the contractual arrangements for increased mammography services 3. Assure that mammography and no-show rates continue to be reported (and remain stable or on upward trend)</td>
</tr>
</tbody>
</table>

**Program: Long live love** ([29–32])

**Setting: School-based**

<table>
<thead>
<tr>
<th>Target: role</th>
<th>Implementation outcomes</th>
<th>Performance objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher: Implementer</td>
<td>1. Teachers reflect and improve on their implementation behavior regarding sexual, reproductive health (SRH) lessons 2. Teachers deliver LLL to students completely 3. Teachers deliver LLL to students according to the guidelines in the teacher manual (fidelity) 4. Teachers deal adequately with the most common difficulties that arise during implementation of SRH</td>
<td>1.1. Teachers reflect critically on their implementation behavior regarding SRH 1.2 Teachers self-monitor and improve the weaknesses in the implementation behavior regarding SRH 2.1. Teachers cover all 6 lessons of LLL (completeness=80% of program) 2.2. Teachers us all program materials of LLL in each lesson 2.3. Teachers cover the most important components of each lesson, as indicated in the teacher manual 3.1. Teachers read the teacher manual as preparation for each lesson 3.2. Teachers deliver each LLL lesson to students according to the teacher manual 4.1. Teachers create a safe and trusted atmosphere in the classroom 4.2. Teachers teach all themes in LLL without shame or taboos interfering with the quality of the lessons 4.3. Teachers handle personal questions of students addressed to themselves depending on their personal need to answer these questions</td>
</tr>
</tbody>
</table>
TABLE 2 | Partial matrices of change objectives for selected examples.

Program: Peace of mind [23, 28]

Behavioral outcome: Patient navigator will complete PMP telephone counseling with eligible patients and complete Appointment reminder calls

<table>
<thead>
<tr>
<th>Performance objectives</th>
<th>Determinants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Awareness and perceptions of PMP</td>
</tr>
<tr>
<td>Patient Navigator searches schedule for upcoming appointments</td>
<td>AP.1.1. Describe requirements of the PMP intervention</td>
</tr>
<tr>
<td></td>
<td>AP.1.2. Describe clinic data system</td>
</tr>
<tr>
<td></td>
<td>AP.1.3. Describe protections for patient information</td>
</tr>
</tbody>
</table>

Program: Long Live Love [29–32]

Behavioral Outcome: Teachers Deal Adequately With The Most Common Difficulties That Arise During Implementation Of Srh

<table>
<thead>
<tr>
<th>Performance objectives</th>
<th>Determinants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Attitude</td>
</tr>
<tr>
<td>1. The teacher integrates the theme of homosexuality as self-evident during all lessons of Long Live Love</td>
<td>A 1.1 Express the importance of a positive attitude of a teacher toward homosexuality during the application of the lessons</td>
</tr>
<tr>
<td>2. Teachers intervene on Homo-negative behavior of students</td>
<td>A 2.3 Express the importance of taking timely measures when students act homo-negatively in the classroom.</td>
</tr>
</tbody>
</table>

not only help identify performance objectives but also the factors influencing whether or not these actions are carried out (determinants). In this way, Implementation Mapping tasks 1 and 2 are iterative. Through the assessment, planners may hear directly from adopters and implementers about the steps required within their setting to achieve the outcomes. Subsequently during task 2, planners may validate the performance objective with the key actors in the implementation setting.

Next, planners identify personal determinants for adopters and implementers. Determinants answer the question of “why?” Why would an implementer deliver the program as planned? (39–41). The barriers and facilitators to implementation are also determinants. Some of these determinants can also be found in the implementation science frameworks or can be theoretical constructs from health promotion theories such as the Social Cognitive Theory (39), Theory of Planned Behavior/Reasoned Action Approach (40), or the Health Belief Model (41). Essentially, determinants are modifiable factors internal to the adopters and implementers that influence their adoption and implementation behavior (13). They are the cognitive reasons why an individual would perform the desired behavioral outcome (in this case an implementation task). For example, outcome expectations, a construct from Social Cognitive Theory (also present in Theoretical Domains Framework), can influence adoption decisions. If a clinic administrator has positive outcome expectations that an evidence-based intervention will increase vaccination uptake within her clinic, she may choose to adopt the program. Alternatively, if she has negative outcome expectations or does not expect the vaccination rate in her clinic to change much due to the evidence-based intervention, she may not adopt the program. Again, Implementation Mapping task one informs this stage of the process as the determinants are often identified through the needs assessment.

Planners then create matrices of change objectives, shown in Table 2 (25). Matrices cross performance objectives with personal determinants to produce change objectives. They answer the question: What has to change in this determinant in order to bring about the performance objective? Change objectives are the discrete changes required in each relevant determinant that will influence achievement of the performance objective. In Table 2, the first performance objective for Peace of Mind is for the Patient Navigator to search the schedule for
appointments, and the relevant determinant is awareness of the Peace of Mind program. These change objectives become the blueprint for developing (or selecting) implementation methods and strategies.

**Task 3. Select Theoretical Methods and Design Implementation Strategies**

In Task 3, planners choose theory- or evidence-based methods to influence the determinants identified in Task 2. They also select or design implementation strategies to operationalize those methods.

Theory-based methods include techniques to influence determinants of implementation (13). These methods can focus on either the individual level (the knowledge, attitudes, and skills of the implementer), or at the organizational level aimed at influencing organizational change directly (e.g., creating institutional commitment and strong organizational leadership). For example, a planner may need to employ information, consciousness raising, persuasive communication, and modeling (theoretical methods) to increase knowledge, address attitudes, and influence outcome expectations (determinants) among potential program adopters (13, 42). Parcel et al. (43) indicate the importance of organizational change for implementation of health promotion interventions, with school health as example. They identify a number of relevant organizational level methods for change, among others: Institutional commitment and strong organizational leadership, primarily from the superintendent, and technical assistance and resources regarding the health promotion intervention. To influence the organizational level, ten Hoor et al. (44) applied the method institutional commitment and strong leadership to the implementation of their strength-based physical exercise intervention. Regular meetings with school managements guaranteed proper participation from the schools and improvement of the study. Multiple methods may be necessary to adequately address a single determinant, and methods often influence more than one determinant. Bartholomew et al. (13) and Kok et al. (42) provide a taxonomy of theory-based methods applicable at the individual- and organizational-levels. Specific methods from the taxonomy relevant to program adoption, implementation, and maintenance include those to increase knowledge; change awareness and risk perception; change attitudes, beliefs, and outcome expectations; change social influence; increase skills, capability, and self-efficacy; change environmental conditions; change social norms and social support; and change organizations, communities, and policies (13). Table 3 provides an example of selected methods and strategies from implementation of the Long Live Love program. A key feature included in this table is consideration of “parameters” of methods used. Parameters represent the guidelines or conditions necessary for a particular change method to be effective. For example, for modeling to be effective, the behavior (of the model) must be reinforced. Decision makers (e.g., clinical medical directors) may not decide to implement a new program simply because a medical director (with whom they identify) has done so. They also must observe that her implementation behaviors were reinforced.

Next, planners select or design implementation strategies to operationalize methods (readers familiar with Intervention Mapping may recall that the operationalization of methods are referred to as practical applications). As previously mentioned, we use the term implementation strategies in Implementation Mapping to refer to both the small-scale strategies to influence specific determinants and change objectives and to the overall package of strategies influencing adoption, implementation, and maintenance behaviors. For example, a fact sheet with heat maps outlining high risk areas is an example of a discrete strategy aimed at increasing knowledge about a health problem among potential program adopters (45). Alternatively, a face-to-face training accompanied by an instruction manual and call
script is an example of a multi component implementation strategy to increase knowledge, self-efficacy, and skills for program implementers (13, 46). While the process we describe here lends itself to developing implementation strategies that match the determinants of implementation behavior, we can also use this method to select strategies that have been used elsewhere.

Table 4 (47) includes information from previous tasks organized into a single table by stage (adoption, implementation, or maintenance), agents, determinants, and change objectives. For example, the Peace of Mind program uses role modeling in a webinar to increase clinic decision makers’ skills and self-efficacy to adopt the program.

In the implementation science literature, Powell et al. (7) identified 73 implementation strategies that can be used in isolation or combination in implementation research and practice. Although comprehensive lists of implementation strategies and their definitions such as these are very important and useful, there is currently little guidance in the implementation science literature about how to select among these strategies to address determinants of implementation. Thus, in practice, the selection (or development) of strategies does not always logically follow from determinants identified. Using Task 3 in Implementation Mapping allows the planner to make decisions about strategy selection or development that logically follow the previous Implementation Mapping steps. The starting point for selection of strategies should always be their suitability to adequately address the determinants. Intervention Mapping draws upon a large body of evidence regarding which methods fit which determinants (42, 48). Additionally, it is very important that the methods are translated into a practical strategy in a way that preserves the parameters for effectiveness and fits with the target population, culture, and context (42). For example, a parameter for role modeling is that the role model needs to show coping or overcoming a barrier, rather than already mastering a skill. By adhering to the parameters of methods, strategies will be more effective for influencing implementation. To have an implementation strategy/implementation intervention to successfully implement a specific intervention/health program.

Figure 2 illustrates how implementation strategies influence health outcomes through their impact on the determinants and behaviors of those responsible for program adoption and implementation and its influence on the implementation context. Similar to logic models of the health promotion program developed using Intervention Mapping, this figure illustrates how implementation strategies can influence the determinants of implementation behaviors (detailed as performance objectives).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Agent</th>
<th>Determinants/change objectives</th>
<th>Theoretical change methods</th>
<th>Practical applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adoption</td>
<td>Clinic decision maker</td>
<td>Awareness/perceptions of PMP&lt;br&gt;Positive attitudes about the innovation – has a relative advantage, not overly complex&lt;br&gt;(from CFIR construct: characteristics of the innovation)&lt;br&gt;Outcome Expectations&lt;br&gt;Skills and Self-efficacy&lt;br&gt;Feedback and reinforcement</td>
<td>PMP program information&lt;br&gt;Persuasion&lt;br&gt;Role modeling</td>
<td>Email blast to BHC members with PMP informational video and link to pre-adoption survey&lt;br&gt;Webinar to BHC members covering evidence-based approaches to breast cancer prevention, PMP information and adoption steps&lt;br&gt;Adoption meeting held with interested clinics&lt;br&gt;Financial assistance to clinics&lt;br&gt;Assistance with connecting to mobile providers to increase screening (as needed)</td>
</tr>
<tr>
<td>Implementation</td>
<td>All</td>
<td>Awareness/perceptions&lt;br&gt;Outcome Expectations&lt;br&gt;Skills and Self-efficacy&lt;br&gt;Feedback and Reinforcement</td>
<td>Cue to participate&lt;br&gt;Communication&lt;br&gt;Mobilization&lt;br&gt;Organizational Consultation/Planning</td>
<td>Invite clinic staff to participate in stakeholder group (templates for invitation email)&lt;br&gt;Email template for site visit (including requested participants) and site visit questionnaire&lt;br&gt;Site visit planning meeting&lt;br&gt;Program implementation guide, clinic handbook, stakeholder manual &amp; computer assisted PMP scripts reviewed during participatory stakeholder meetings&lt;br&gt;Implementation readiness checklist&lt;br&gt;Stakeholder meetings to support implementation (continue after reminder calls begin). E-newsletter shared with stakeholders</td>
</tr>
<tr>
<td>Implementation</td>
<td>Program champion navigator</td>
<td>Awareness/Perceptions&lt;br&gt;Outcome Expectations&lt;br&gt;Skills and Self-efficacy&lt;br&gt;Feedback and Reinforcement</td>
<td>Information&lt;br&gt;Persuasion&lt;br&gt;Skill building and guided practice&lt;br&gt;Modeling&lt;br&gt;Monitoring and feedback&lt;br&gt;Technical assistance/capacity building&lt;br&gt;Facilitation&lt;br&gt;Vicarious reinforcement</td>
<td>Face to face training held over two 4 h sessions. Training was submitted to Texas for CEU certification for community health workers and social workers&lt;br&gt;BHC navigators model EBI behavior and provide ongoing implementation support on-site&lt;br&gt;PMP research team available via email, phone and training booster sessions as needed&lt;br&gt;Paperwork processes to provide funds for patients needing financial assistance from PMP</td>
</tr>
</tbody>
</table>
for adoption, implementation, and maintenance) which in turn influence implementation outcomes.

**Task 4. Produce Implementation Protocols and Materials**

The next task in Implementation Mapping is to produce implementation protocols, activities and/or materials. Similar to Step 4 in Intervention Mapping, this requires planners to create design documents, draft content, pretest and refine content, and produce final materials. Even when selecting already existing strategies [e.g., from the Expert Recommendations for Implementing Change (ERIC) list] the content within these strategies must be defined. Using Implementation Mapping, it is clear what messages, methods, and materials are needed rather than simply having selected a general strategy. Design documents are shared between planners and production teams, and they are created for each document or other materials that are a part of the implementation strategy. While no two design documents will be the same, they may include the following types of information: purpose of the material, intended audience, targeted determinants and change objectives, theoretical methods, draft content, a description of appropriate imagery, or a flowchart. For example, a planner might want to produce a testimonial video highlighting program successes in the community. This video will be posted on the program’s website and target future adopters. A design document from the planners may include the following: (1) the overall purpose of the video; (2) a description of the potential adopters; (3) determinants such as knowledge, outcome expectations, and perceived social norms and associated change objectives; (4) a list of the relevant theoretical methods such as modeling, persuasive communication, and information; and (5) draft interview questions to ask the video’s subject. This document provides the production team with all of the information necessary to conduct an interview and produce the testimonial video. These design documents do not only support the development of implementation interventions, but can also help evaluation and potential adaptation of implementation interventions.

**Task 5. Evaluate Implementation Outcomes**

Interventions cannot be effective if they are not implemented, and their effectiveness will be compromised if they are implemented incompletely. Therefore, implementation outcomes are essential preconditions for achieving desired changes in behavior, health, or quality of life outcomes (49). Following Implementation Mapping tasks 1–4 increases...
the likelihood of developing implementation strategies that address identified barriers and enable implementation. Nevertheless, it is essential to evaluate whether or not these strategies have led to intended adoption, implementation, and sustainability outcomes.

Understanding implementation generates information to improve the intervention and its delivery, and for interpreting its effects on intended outcomes. Implementation evaluation and process evaluation are terms that are often used interchangeably and essentially assess the extent to which implementation strategies fit well within the context, are delivered with fidelity and are addressing identified needs (50, 51). Process and implementation evaluation can answer questions such as who the program reached, to what extent was it delivered as planned (to whom, what level of fidelity, whether theory, and evidence-based change methods were applied correctly). Because implementation is highly dependent on context, process evaluation questions can also include those that assess the organizational factors that influenced intervention adoption, use, and/or maintenance including understanding what were the barriers and facilitators to implementation.

Procter and colleagues defined several types of implementation outcomes including acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability (49). In this task we describe how to use the preceding tasks to develop a plan to evaluate implementation and determine the impact of implementation strategies developed following tasks 1–4.

Analogous to Step 6 (Evaluation Plan) of Intervention Mapping, this task (Task 5) in Implementation Mapping helps the planner write effective process evaluation questions, develop indicators and measures for assessment, and specify the process/implementation evaluation design. Using Implementation Mapping, the planner describes expected implementation outcomes (for adoption, implementation, and/or maintenance) and performance objectives. The performance objectives delineate the specific implementation actions needed to deliver the intervention. These can be used to develop instruments to assess fidelity. Likewise the identification of determinants of implementation and creation of matrices of change objectives that state the needed changes in determinants to produce implementation outcomes, help identify important potential mediators or moderators of implementation outcomes and can again be used to develop measures to detect change in those mediators or moderators.

Following identification of process evaluation questions and measures, it is important to consider potential designs for assessing implementation outcomes. Efforts to implement evidence-based interventions are often complex, employ multilevel implementation strategies, and involve different stakeholders. The use of mixed methods approaches is particularly useful for evaluating implementation outcomes (52, 53); quantitative approaches can help confirm hypothesized relationships between implementation strategies, their impact on determinants, and the subsequent impact on implementation outcomes, while qualitative methods can explore important contextual factors influencing these relations and obtain deeper and more nuanced information about reasons for successes and failures (54). Palinkas et al. (54) provide recommendations for mixed methods approaches including the use of purposeful sampling in mixed methods implementation research.

A critical perspective in the use of Implementation Mapping for planning and evaluating implementation strategies is that, like Intervention Mapping, it is an iterative endeavor. It is unlikely, for example, for the needs and asset assessment (Task 1) to identify all barriers and facilitators to implementation and that these will likely emerge during the planning process, particularly when choosing appropriate applications of change strategies to influence determinants. Likewise, during process evaluation, it may be obvious that some key determinant was missed or that the delivery approach is not maximizing reach. The framework allows for planners to cycle back to previous tasks to more accurately reflect the mechanisms influencing implementation as well as make changes to the strategies to maximize impact.

**Implementation Logic Model**

The products of Tasks 1–5 of Implementation Mapping can be presented in a model that illustrates the logic of how the strategies will affect implementation and effectiveness outcomes (see Figure 2). The logic goes from left to right with the innovation (intervention, program, policy, practice) on the far left followed by implementation strategies that deliver methods that influence determinants that change implementation behaviors and conditions and lead to implementation and ultimately effectiveness outcomes. The planning process, however, goes from right to left beginning by articulating desired outcomes and the adoption, implementation, and maintenance behaviors and conditions that will bring about those outcomes, then describing the determinants that lead to those behaviors and conditions, and finally selecting methods and developing strategies that will ultimately bring about desired outcomes. The logic model created as part of the process for planning or selecting implementation strategies helps describe the mechanisms through which we expect the implementation strategies to work. This, together with the matrices of change objectives produced in Task 2, represent blueprint or maps for the implementation strategies and guide decisions along the development or selection process. The implementation logic model is useful for both planning the implementation strategies and for designing their evaluation.

**Using Implementation Models to Inform Implementation Mapping**

The Implementation Mapping process provides a framework for using implementation models for planning or selecting implementation strategies. For example, the Interactive Systems Framework (ISF) (55) can help identify key actors including adopters and implementers within particular settings. Reach, Effectiveness, Adoption, Implementation, and Maintenance, or RE-AIM, may help implementation strategy planners organize implementation outcomes at multiple levels including individual and organizational levels (56). Additionally, the Consolidated Framework for Implementation Research (CFIR) can help guide decisions about contextual factors that may influence program adoption and implementation (57). This can inform
the development of performance objectives or determinants that will enter into the matrices constructed in Task 2. Combined, these models can be used to develop implementation strategies that take into account specific contexts for program adoption and implementation.

For example, they can help the planner identify program targets that go beyond effectiveness outcomes and consider adoption, implementation and maintenance (e.g., RE-AIM). In Task 1, they can be used to inform who the adopters and implementers may be and in Task 2, describe the necessary actions to adopt or deliver a program, practice, or policy.

The CFIR (57) and the Interactive Systems Framework (ISF) can help planners identify contextual and motivational factors relevant to program adoption and implementation. They can also help identify the types of capacity building that may be required to enhance implementation. The CFIR describes constructs related to implementation including (perceived) interventions characteristics (e.g., the source of the intervention), outer setting (e.g., patient needs and resources), inner setting (e.g., the implementation climate), individuals’ characteristics (e.g., self-efficacy and the implementation process (e.g., opinion leaders) (57, 58). It can therefore be useful when identifying individuals involved in implementation or in control of certain contextual factors (Task 1) and can also help identify actions needed to change implementation behaviors or contexts and their determinants (Task 2). For example, in studies implementing a Chronic Care Model (CCM) in primary care settings, implementation facilitators included a number of CFIR constructs such as engaged leadership, positive beliefs about the model, networks and communication, organizational culture, implementation climate, and structural characteristics of the setting (59). Barriers included lack of leadership engagement, lack of readiness for implementation, and poor execution (59). Therefore, researchers seeking to implement CCM in additional primary care settings and aiming to plan or select strategies could use Implementation Mapping informed by CFIR to describe performance objectives related to engaging leadership, building enthusiasm for CCM, and then identifying the determinants influencing these actions. IM would then help identify methods and plan or select strategies to address those determinants.

Another specific example of how CFIR may inform the Implementation Mapping process is as follows: if the CFIR construct leadership engagement is found to be an important predictor of implementation this can help create performance objectives (created by asking: What does the leader have to do to increase engagement?) as well as determinants (Why would they engage?). Likewise, CFIR constructs related to perceptions of the innovation (e.g., relative advantage) can point to potential determinants of both adoption and implementation behaviors. Table 4 includes CFIR informed determinants (relative advantage and complexity) and how they fit in the mapping process.

The ISF and its Readiness concepts (Readiness = Motivation × Innovation Specific Capacity × General Capacity- R = MC²) (55) can help identify determinants of (Task 2) and methods (Task 3) for enhancing adopters’ and implementers’ readiness for implementation. Further, using the ISF, planners can think through the process of adoption and implementation at multiple levels and identify key actors at each of them (60). Another framework that is often used to understand determinants of behavior and guide implementation is the Theoretical Domains Framework (TDF) (55). The TDF includes 84 constructs listed under one of the 14 domains, all derived from the 83 theories of behavior and behavior change identified (61). It has been used to identify barriers of HPV-related clinical behaviors for general practitioners and practice nurses (62); to understand anesthesiologists’ and surgeons’ routine pre-operative testing behavior in low-risk patients (63); and to understand treatment adherence of adults with cystic fibrosis (64). Note that TDF constructs can be determinants (related to the wanted behavior—such as self-efficacy) as well as methods (e.g., goal setting). In the systematic and iterative process of Implementation Mapping, these belong to task 2 and task 3, respectively.

Thus, informed by these frameworks and guided by the implementation mapping protocol, program planners can carefully select key implementers, articulate implementation behaviors and determinants, and then select methods and strategies to address them.

Examples
One example of the application of Implementation Mapping is the development of strategies to implement the “Focus on Strength” program by ten Hoor et al. (65, 66). The Focus on Strength program is a school-based physical activity intervention that included 30% additional strength exercises in the physical education classes (about 15 min per session, 3 times per week) to especially reach overweight children who may be less fit but stronger than their classmates, thus allowing them to have some success and build self-efficacy. Additionally, teachers gave monthly motivational lessons to promote autonomous motivation of students to become more physically active outside school. In task 1, the planners identified adopters and implementers: managers and teachers. In task 2, they identified adoption and implementation outcomes, and their determinants. While the addition of the extra lessons seemed necessary, this was very difficult to implement in schools with already time-constrained curricula. After consulting the implementers (particularly the physical education teachers, but also the managers and planners), “time” was identified as an important potential barrier. In task 3, the planners chose methods (such as participatory problem solving and technical assistance) and strategies (teacher workshops and a workbook). To facilitate implementation, they decided (together with the implementers) to limit the extra strength component in the physical education lessons to 30% of the physical education time (about 15 min per lesson or 45 min per week) and 1 motivational lesson per month (about 10 lessons per year). This improved feasibility and facilitated adoption, implementation, and maintenance of the program. In task 4, the planning group developed and successfully used teaching protocols and materials. This way, understanding the implementation setting, including key actors (e.g., curriculum planners, directors, and teachers) and potential barriers and facilitators to adoption and
Implementation, potential reasons not to adopt or implement the intervention can be overcome.

Another example, described in a recently published study, also used Implementation Mapping (Intervention Mapping Step 5) to plan an implementation intervention to increase adoption, implementation and maintenance of the Peace of Mind Program, an intervention to increase mammography screening among patients of federally qualified health centers (FQHCs) (25). The authors describe how the planning group, including stakeholders, participated in brainstorming and discussing answers to questions posed by each of the Implementation Mapping tasks. They identified clinic leaders as adopters and mammography program staff and patient navigators as implementers. They then identified performance objectives and determinants based on feedback from stakeholders and using the CFIR (57, 58) “process of implementation” and “inner setting” domains to help inform the identification of both motivational and contextual factors influencing participation. This helped in the identification of performance objective and determinants. They then matched theoretical methods with determinant and operationalized them as strategies (25).

**DISCUSSION**

Despite significant advances in clinical, health promotion, and policy research that produce effective intervention, the gap between research and practice limits their impact on improving population health (3, 67). Closing this research to practice gap requires powerful strategies to address the multi-level barriers and facilitators to adoption, implementation, and maintenance needed to accelerate and improve delivery of evidence-based interventions. The study of and use of implementation strategies is central to the National Institutes of Health’s (NIH’s) mission of increasing the impact of the nation’s investment in health-related research (68). Implementation science literature demonstrates a growing body of work on dissemination and implementation models and frameworks in the last several years (8, 9). These frameworks describe determinants, systems and processes necessary for active dissemination and implementation as well as implementation outcomes; yet, they leave some gaps in procedural knowledge on how to use these frameworks to inform the development of effective implementation strategies. As a result, few studies use theory in developing implementation strategies and sometimes researchers are not aware of the evidence-base of the methods they employ (14). This paper described a detailed systematic process for developing implementation strategies that is informed by theory, evidence, and participatory approaches to planning. Through the development of logic models, Implementation Mapping can also help better define and understand the mechanisms through which implementation strategies lead to desired outcomes.

Despite efforts to better classify implementation strategies (7) and better articulate who enacted the strategy, its influence on determinants, and its effectiveness (69), confusion remains about how to develop them and what the mechanisms of action may be. There has been much confusion in the field, for example, related to a failure to distinguish between mechanistic (theoretical methods or techniques) that cause changes in behavior, and how they are operationalized in the practice or community setting (strategy). For example, Ivers et al. (70) state that the use of audit and feedback “is based on the belief that healthcare professionals are prompted to modify their practice when given performance feedback showing that their clinical practice is inconsistent with a desirable target.” That is correct, but it is a method which has proven effectiveness and stems from theories such as Theories of Learning, Goal-setting Theory and Social Cognitive Theory (13, 42). Audit and feedback is indeed a frequently used method with a strong theoretical underpinning. It is also one of the “strategies” listed in the refined ERIC. However, the refined ERIC does not refer to the theoretical bases of their listed strategies, and the strategies listed are often broad recommendations (e.g., develop health education material) or guidelines. The different ways constructs, methods, strategies, etc., are classified across various compilations and frameworks gives room for confusion and misunderstanding. In this paper, we propose an organizing and conceptual framework to develop or select strategies that are specifically mapped to identified determinants of implementation and contain change methods powerful enough to address them.

An important contribution that Implementation Mapping can make to the field of implementation science literature is in filling the conceptual and practical gap between identifying implementation barriers and facilitators and developing or selecting implementation strategies. Without this type of systematic guidance for the development of implementation interventions, we will continue to struggle as a field in both the development and the selection of theory and evidence-based implementation strategies most likely to influence change.

Recently, authors have highlighted the need to articulate the causal pathways through which implementation strategies are effective (71). They suggest the need to link strategies to barriers and describe not only the desired proximal and distal outcomes but also the processes or mechanisms through which implementation strategies are effective (71). A foundational principle of Intervention Mapping and Implementation Mapping is the development of logic models (causal models) that illustrate the causal pathway between the implementation strategy, the methods it operationalizes (mechanisms), the determinants of implementation affected and the proximal and distal implementation outcomes. This includes changes in implementation behavioral and contextual factors, implementation outcomes, and the ultimate impact on health and quality of life.

Another recent article (72) suggests a process for creating a tailored implementation blueprint that includes identification of determinants of implementation. This suggestion is analogous to our Task 1 of Implementation Mapping and the selection or matching of strategies (as in Implementation Mapping Tasks 3 and 4). The importance of planning implementation strategies using a collaborative process including stakeholders at multiple levels is another central element of Implementation Mapping as described above. A recent example of one way to do this is conjoint analysis (72). We agree with recent
recognition of the pressing need for processes to select and match strategies that fit implementation needs and contexts and believe that Implementation Mapping is a potential solution (17). Intervention Mapping for Planning Implementation Strategies, what we have called Implementation Mapping here, has already been employed by several authors (25, 26, 73–75) and recommended as an effective approach (17).

Implementation Mapping can advance the field of Implementation science by (1) elucidating mechanisms of change (i.e., how implementation strategies influence outcomes through change in implementation determinants) (2) better guiding the use of implementation models and frameworks during the planning process, and (3) improving the impact of implementation strategies on outcomes. The use of logic models of change that delineate the hypothesized relationships between causal factors (implementation barriers, contextual factors, behavior, and organizational change methods) and implementation outcomes can guide the development and selection of implementation strategies that will have the greatest potential impact on implementation and health outcomes.

Future directions include studies to better understand how existing implementation frameworks and models can inform the planning process. Although we believe that Implementation Mapping can help, we are only beginning to describe and demonstrate the best ways that implementation frameworks and models (and the constructs within them) can inform the development and selection of implementation strategies. Answers to questions about which tasks of Implementation Mapping are best informed by which models or elements of models is still evolving. Additionally, studies to explicitly test the use of Implementation Mapping as a planning framework for implementation strategies as compared to other methods can help provide evidence of the utility of the process.

CONCLUSION

Too many evidence-based interventions are not put into practice, or are eventually implemented but with a significant delay. This compromises the potential of research findings in improving health care and health promotion efforts, and subsequently health outcomes. Implementation Mapping outlines a practical method for planning implementation strategies that will be optimally effective. Just as the systematic planning of health promotion and other interventions have greatly improved their effectiveness, the use of Implementation Mapping to plan implementation strategies will improve the appropriateness, quality, and impact of these strategies on implementation outcomes. Consequently this will lead to increased adoption, implementation, and sustainment of evidence based interventions and overall improvement in population health.

AUTHOR CONTRIBUTIONS

MF, GH, SvL, RB, and SR: wrote sections of the manuscript. GP, RR, CM, and GK: participated in the development of implementation mapping process, reviewed and edited the manuscript.

FUNDING

SR was supported by a predoctoral fellowship from the University of Texas School of Public Health Cancer Education and Career Development Program—National Cancer Institute/National Institutes of Health Grant R25 CA57712. MF was partially supported by the National Cancer Institute 1 R01/CA 163526 and the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services contract/grant number R18HS023255-03.

ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of our late colleague, L. Kay Bartholomew Eldredge, the lead creator of Intervention Mapping. We also wish to acknowledge the work of other colleagues such as Nell Gottlieb, Patricia Dolan Mullen, Melissa Peskin, Belinda Hernandez, Linda Highfield, Andrew Springer, Melissa Valerio, Lara Staub Savas, Cam Escoffrey, Byron Powell, and others who have applied Intervention Mapping for the adaption and implementation of interventions and whose work and suggestions have helped improve the process. We also wish to thank Kelly McGauhey and Marsha Lee for editorial assistance.

REFERENCES

Implementation Climate and Time Predict Intensity of Supervision Content Related to Evidence Based Treatment

Michael D. Pullmann 1*, Leah Lucid 2, Julie P. Harrison 2, Prema Martin 2, Esther Deblinger 3, Katherine S. Benjamin 2 and Shannon Dorsey 2

1 Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, United States, 2 Department of Psychology, University of Washington, Seattle, WA, United States, 3 CARES Institute, Rowan University School of Osteopathic Medicine, Stratford, NJ, United States

Objective: Children infrequently receive evidence-based treatments (EBTs) for mental health problems due to a science-to-practice implementation gap. Workplace-based clinical supervision, in which supervisors provide oversight, feedback, and training on clinical practice, may be a method to support EBT implementation. Our prior research suggests that the intensity of supervisory focus on EBT (i.e., thoroughness of coverage) during workplace-based supervision varies. This study explores predictors of supervisory EBT intensity.

Methods: Participants were twenty-eight supervisors and 70 clinician supervisees. They completed a baseline survey, and audio recorded supervision sessions over 1 year. Four hundred and thirty eight recordings were coded for supervision content. We chose to explore predictors of two EBT content elements due to their strong evidence for effectiveness and sufficient variance to permit testing. These included a treatment technique (“exposure”) and a method to structure treatment (“assessment”). We also explored predictors of non-EBT content (“other topics”). Mixed-effects models explored predictors at organizational/supervisor, clinician, and session levels.

Results: Positive implementation climate predicted greater intensity of EBT content coverage for assessment (coefficient = 0.82, p = 0.004) and exposure (coefficient = 0.87, p = 0.001). Intensity of exposure coverage was also predicted by more time spent discussing each case (coefficient = 0.04, p < 0.001). Predictors of greater non-EBT content coverage included longer duration of supervision sessions (coefficient = 0.05, p < 0.001) and lower levels of supervisor EBT knowledge (coefficient = −0.17, p = 0.013). No other supervisor- or clinician-level variables were significant predictors in the mixed effects models.

Conclusion: This was the first study to explore multi-level predictors of objectively coded workplace-based supervision content. Results suggest that organizations that expect, support and reward EBT are more likely to have greater intensity of EBT supervision coverage, which in turn may positively impact clinician EBT fidelity and client outcomes.
INTRODUCTION

Many evidence-based treatments (EBTs) have been developed to address child and adolescent mental health needs (1). However, the potential promise of EBTs has not been realized due to the substantial challenge of implementing them in community mental health settings (2–4). Growing consensus in the literature indicates that EBTs are implemented at a slow pace in community settings, leading to critical gaps in the quality and effectiveness of mental health care (5–7). Experts have categorized over 70 implementation strategies (8), one of which is providing clinical supervision. Generally, clinical supervision is defined as an evaluative intervention wherein more senior clinicians provide oversight to more junior clinicians in order to ensure the quality of their services and provide ongoing clinical training (9). In the Exploration, Adoption/Preparation, Implementation, and Sustainment model of EBT implementation (EPIS) (10), fidelity monitoring and support—important aspects of EBT-focused clinical supervision—are specifically noted as inner setting factors affecting implementation. Without ongoing clinical supervision focused on the EBT, clinicians’ fidelity can be low (11, 12), creating challenges for both the active implementation and sustainment phases of EBT implementation (10).

Clinical supervision focused on EBT delivery has been demonstrated to improve clinician EBT fidelity (13), knowledge, attitudes and skills (14). In relevant work from the expert consultation literature, in which EBT-specific supervision was provided by external EBT experts, a greater dose of EBT-focused supervision resulted in greater clinician skill in the EBT (15). Active learning strategies used in supervision (e.g., modeling) predicted community mental health clinicians’ competent use of EBT strategies in the next therapy session (16). In an analog study that randomized psychology trainees into two groups (supervision as usual vs. supervision with active learning elements), only the active learning group had greater clinician knowledge, attitudes, and skill (14).

Workplace-Based Supervision in Community Mental Health Organizations

One potentially sustainable way to increase clinician receipt of EBT-focused supervision in community settings with limited resources to support ongoing expert consultation is to identify existing organizational supports in which to embed EBT coverage. In a national survey, most community mental health organizations reported providing weekly workplace-based clinical supervision (17). Workplace-based supervision includes both clinical supervision as well as oversight for administrative issues, professional development, and emotional support, provided by internal staff employed within an organization (18). In a study by our research group examining workplace-based supervision within organizations participating in a state-funded EBT initiative, weekly occurrence of supervision was mostly upheld [75% reported weekly supervision for ∼1 h (19)].

Workplace-Based Supervision and EBT Implementation

Very limited research has focused on workplace-based supervision and EBT implementation (20, 21). In one study examining discussion of evidence-based principles for behavior disorders, clinicians, and supervisors reported that EBT coverage was generally brief (20). In a study focused on Trauma-focused Cognitive Behavioral Therapy (TF-CBT) implementation (22), clinicians reported moderate coverage of TF-CBT elements in supervision (23). Schoenwald and colleagues (21) trained workplace-based supervisors in a manualized supervision model designed to support the implementation of Multisystemic Therapy (24). Supervisor adherence to treatment principles was related to increased treatment fidelity, and supervision structure was related to speed of change in client symptoms and functioning.

In a study on which the current investigation builds, Dorsey et al. (25) objectively coded the workplace-based supervision sessions of supervisors and clinicians participating in a state-funded TF-CBT initiative. TF-CBT is an evidence-based treatment for mental health sequelae subsequent to trauma exposure (26). It includes nine treatment elements (22): psychoeducation, parenting, relaxation, affect modulation, cognitive coping, trauma narrative, and processing trauma-related thoughts (imaginal exposure: facing up to memories of the traumatic event), in vivo mastery of trauma reminders (situational exposure: facing up to reminders in the environment), conjoint sessions, and enhancing safety. Many of these are used in other cognitive behavioral approaches to child and mental health disorders. Sixteen content areas, described in a measures table in Appendix A in Supplementary Material, were coded for occurrence and intensity of occurrence in the Dorsey et al. (25) study. These content areas included the nine TF-CBT elements as described in Table 1, some of which
were collapsed for coding feasibility, as well as other content necessary for supervising TF-CBT (i.e., assessment, child's trauma history, use of art, play and books to engage children, treatment engagement), and clinician-level EBT techniques found to be infrequently used by clinicians in usual care (5) (e.g., assigning/reviewing client homework), and essential for effective delivery of TF-CBT [see Appendix A in Supplementary Material and (25) for more information on coding procedures]. There was substantial variation in content coverage, with some elements covered in more than half of the supervision sessions, and other important elements covered more rarely.

Potentially more important than whether an element is covered, is the intensity with which a supervisor covers EBT content. Following McLeod and Weisz's (27) operationalization, we define intensity as the frequency and thoroughness with which specific content elements are covered. As an example, the EBT of focus for this study, TF-CBT, includes two content elements focused on exposure (see Table 1, the trauma narrative imaginal exposure and in vivo exposure). These two exposure content elements were collapsed for coding of exposure content coverage in supervision sessions in the Dorsey et al. (25) study, on which this study builds [see Appendix A in Supplementary Material and Dorsey et al. (25) for more details about the coding procedures]. Exposure gradually reduces anxiety by having clients repeatedly face a feared stimulus, such as memories of a traumatic event. Intensity of exposure coverage during supervision would be determined by the extent of detail and time spent planning exposure content for an upcoming session or debriefing exposure coverage for a completed session. High intensity coverage of exposure may involve a detailed discussion of exposure use in the last session and planning for the next session (e.g., whether and how caregivers would be involved, ways to support the client during exposure, identifying strategies to manage client avoidance). Low-intensity coverage of exposure involves only a brief mention (e.g., “You should start the trauma narrative”). This low intensity coverage of EBT elements (e.g., a brief mention) is unlikely to provide sufficient fidelity monitoring or support. Similarly, assessment is a commonly used technique in TF-CBT that is discussed in supervision with varying levels of intensity. It is defined as the discussion of information about the client's psychosocial symptoms or behavior problems from standardized, formal assessment measures or functional analysis. Assessment is not one of the nine TF-CBT clinical content items but is necessary for delivering and supervising TF-CBT. High intensity coverage for assessment would involve the supervisor and clinician planning for assessment, reviewing assessment scores, and considering implications of scores for treatment. Low intensity would involve a brief mention of assessment without further discussion and would be unlikely to be related to any modifications in treatment planning or clinical approach.

**Study Purpose and Rationale**

The current study extends Dorsey et al. (25) and seeks to identify clinician, supervisor, and organization characteristics that predict the intensity of coverage for two specific content elements important for workplace-based supervision of EBTs. By identifying the predictors of EBT-focused supervision content, this study can provide valuable information to optimize the effectiveness of clinical supervision as an implementation strategy. We chose to focus this study on the content elements of exposure and assessment for two primary reasons. First, there were statistical limitations that prohibited analyses predicting the variance of many other content elements. In the Dorsey et al. (25) study, both were among a small subgroup of content items that had sufficient overall variance in intensity of coverage and variance at the clinician and/or supervisor levels to permit the investigation of predictors. Second, exposure and assessment were selected from the subgroup because of their theoretical importance for TF-CBT implementation. We also focused on non-EBT related “other topics” content as an analytic counterpoint.

EBTs are generally comprised of multiple clinical intervention elements, but also include structural elements that support and organize technique delivery (28). Exposure was included in this study because it is a common and effective technique used in EBTs for child and Adolescent anxiety disorders (29), is included in almost all EBTs for trauma treatment (19), and is one of the most active ingredients of TF-CBT (see Table 1). Exposure has very strong evidence for effectiveness (30); some studies even found exposure to be just as effective alone as when combined with other active components (31, 32). Despite the robust evidence supporting exposure, clinicians use it only rarely (27), possibly due to lack of comfort and training with this technique (33). Assessment is included in this study because it is a common and effective structural element that supports delivery of any EBT by assisting in planning for which EBT to use and if the client is having symptom improvement with receipt of the EBT. For flexible treatments, including TF-CBT, assessment also assists clinicians in deciding which clinical elements to deliver and when to deliver them. When used as part of routine outcome monitoring (e.g., repeated administration and review), it has been demonstrated to increase quality of care and improve outcomes (34, 35). Regular administration and review of client assessments helps focus clinicians on the needs of their clients and systematically identify progress or lack thereof (36). In TF-CBT, clinicians are expected to assess clients for trauma exposure and mental health symptoms before beginning treatment and to continue to assess clients’ mental health symptoms throughout treatment to guide element ordering, dose (how many sessions allocated to any element), and to determine treatment response (i.e., is the client making progress?). The coverage of assessment in supervision could possibly facilitate thoughtful and timely treatment adjustments. Therefore, we chose to study assessment as a complement to exposure because assessment represents an evidence-based structure that supports treatment, rather than a specific clinical element like exposure.

In addition to examining predictors of two EBT content elements, we also wanted to examine predictors of non-EBT content coverage (i.e., *other topics*). *Other topics* was defined as discussion of issues unrelated to the child’s traumatic experiences or TF-CBT practice components. This content may include the case background information, crisis, or case management, administrative work, and non-work related conversations. With limited time per case (25), the EBT focus of supervision could be...
TABLE 1 | Content elements of Trauma-Focused Cognitive Behavioral Therapy (TF-CBT).

<table>
<thead>
<tr>
<th>Content element</th>
<th>Goal</th>
<th>Description and Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PHASE 1: STABILIZATION AND SKILL BUILDING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychoeducation</td>
<td>Normalize parent and child’s symptoms, provide information about responses to trauma, emphasize</td>
<td>Orienting the parent to the TF-CBT model by explaining the child’s symptoms and the collaborative nature of treatment, and gives hope by describing the researched effectiveness of TF-CBT</td>
</tr>
<tr>
<td>Parenting skills</td>
<td>Improve parental functioning, which is related to child outcomes; structure and predictability enhances adaptive functioning for child and parent</td>
<td>Teaching and practicing functional analysis, praise, selective attention, time out, and contingency reinforcement</td>
</tr>
<tr>
<td>Relaxation skills</td>
<td>Reduce physiological symptomology related to anxiety or trauma</td>
<td>Teaching and practicing focused breathing, mindfulness, meditation, and progressive muscle relaxation</td>
</tr>
<tr>
<td>Affective modulation</td>
<td>Help children voice and handle their feelings more effectively with the goal of reducing avoidant strategies</td>
<td>Teaching and practicing feeling identification, thought interruption and positive interruption, positive self-talk, enhancing problem solving and social skills, managing difficult affective states</td>
</tr>
<tr>
<td>Cognitive coping</td>
<td>Explore thoughts and challenge maladaptive thoughts</td>
<td>Education about the cognitive triangle, and teaching to recognize types of inaccurate or unhelpful thoughts</td>
</tr>
<tr>
<td><strong>PHASE 2: TRAUMA NARRATION AND PROCESSING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trauma narration and processing*</td>
<td>Gradual imaginal exposure to the trauma and surrounding events, thoughts, and feelings to unlink trauma reminders to negative feelings</td>
<td>Taking and writing about the trauma gradually, but in detail, with the help of a therapist</td>
</tr>
<tr>
<td><strong>PHASE 3: CONSOLIDATION AND CLOSURE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In vivo mastery of trauma reminders</td>
<td>Exposure to objects and experiences to unlink trauma reminders to negative feelings. The only piece which is optional because most children do not overgeneralize fear to objectively non-threatening stimuli and so do not require in vivo exposure to combat functionally impairing avoidance</td>
<td>Gradually allow child to adjust to a feared situation that is objectively safe</td>
</tr>
<tr>
<td>Conjoint child-parent sessions</td>
<td>Encourages parents and children to practice skills together and to make the child more comfortable discussing the trauma with the parent</td>
<td>Parent should be carefully prepared to increase likelihood of positive interactions between parent and child in session</td>
</tr>
<tr>
<td>Enhancing future safety and development</td>
<td>Increase the likelihood of personal safety; especially important when there is potential for ongoing trauma</td>
<td>Developing a personal safety plan, teaching related skills: communicating feelings, attending to “gut feelings,” identifying safety cues, learning body ownership, recognizing secrets vs. surprises, and how to ask for help</td>
</tr>
</tbody>
</table>

*In Dorsey et al. (25) and the current study these two elements were collapsed and coded as “exposure” to capture exposure content coverage in supervision.

“crowded out” by coverage of other topics that may be clinically relevant but do not directly support the clinician in the EBT. We included this variable as a negative control outcome (37) and a counterpoint to the other two EBT-focused dependent variables. Therefore, if a variable predicts intensity of both EBT and non-EBT content, we might conclude that it is simply a broad facilitator of intensity of supervision in general. However, if a predictor is positively related to EBT content intensity and also negatively related or unrelated to non-EBT content, it provides some empirical justification that the predictor may be a specific mechanism of intensity of EBT content coverage.

**Potential Predictors of EBT Content Coverage in Supervision**

Because there is limited research on predictors of supervision content, we draw our hypothesized predictors from other supervision-focused research [e.g., (16)], theoretical models of supervision (38, 39), the expert consultation literature (40), and predictors of clinician EBT practice (41). Figure I displays our overall theoretical model and the placement of the current study within that model. Based on the studies described above, our overall theoretical model proposes that supervision acts as an implementation strategy that positively moderates the relationship between EBT training and EBT implementation (adoption, fidelity, and sustainment). The effectiveness of supervision as an implementation strategy is positively moderated by the intensity of EBT-related content delivered during supervision, and negatively moderated by the intensity of non-EBT content. In regards to the tested part of the model, we hypothesized that intensity of coverage for the two EBT-related supervision content areas, exposure and assessment, would be predicted by multiple characteristics of the organization, supervisor, clinician, and session, described in detail below.

**Organizational Factors**

Implementation climate, defined as employees’ shared perceptions of the degree to which innovation use is expected, supported, and rewarded (42), may be an important organizational-level predictor, given its role in theoretical
models of organizational effectiveness (43, 44), though empirical work on implementation climate has been limited (43). If an organization expects, supports, and rewards EBT delivery, we believe supervisors and clinicians would be more motivated to address EBT-related content during supervision. A few cross-sectional studies have tested Klein & Sorra’s model (42) and found support for the relation between implementation climate and implementation effectiveness [e.g., implementation of computer technology in schools (45) and physician enrollment of patients in clinical trials (46)]. In studies focused on workplace-based supervision, our research team has found an association between implementation climate and self-reported greater coverage of clinical versus administrative content (19) and intensity of TF-CBT-specific content (23). Therefore, we hypothesized that implementation climate will be positively associated with intensity of EBT content, and negatively associated with intensity of non-EBT content.

**Supervisor Factors**

As supervision is an interpersonal interaction between the supervisor and the clinician supervisee, individual characteristics likely play a role in determining the nature of the interaction. Our hypotheses are informed by research findings that clinicians’ training, experience, and skill have been associated with client outcomes (47, 48), that clinicians’ years of experience has been associated with client satisfaction (49), and that clinicians’ theoretical orientation has been associated with the use of EBT strategies in treatment (41). For EBT content to be covered during supervision, supervisors must have some expertise with the EBT (measured by their amount of training, whether they primarily use EBTs, and an objective EBT knowledge test), they must have a belief in their own abilities to cover EBT (measured by self-efficacy and self-rated skill), and they must have the willingness to cover EBT (measured by attitudes toward EBTs, CBT theoretical orientation, and comfort with providing supervision on specific EBT elements). Therefore, we hypothesized that these indicators would be positively associated with supervision intensity of EBT content and negatively associated with non-EBT content. We explored for the impact of other supervisor characteristics that although not specific to EBT, may play a role in supervision content coverage, including years of experience conducting therapy, percent of time providing supervision, and their own ongoing involvement in providing therapy.
Clinician Factors
Clinician characteristics may also be associated with coverage intensity, perhaps directly through asking for or steering the supervision session in certain directions, or indirectly through supervisors’ reactions to clinician characteristics. For instance, one study found that supervisors provided more professional development to clinicians whose clients demonstrated weaker improvements, possibly reflecting supervisors’ perceptions of a need for improving clinical skill (21). Similar to supervisors, we felt that EBT content would be impacted by clinicians’ expertise, belief in their own abilities to provide EBTs, and willingness to engage in the content. Therefore, we hypothesized that EBT training, objectively measured knowledge, self-efficacy, self-rated skill, attitudes toward EBTs, and CBT theoretical orientation would be positively associated with intensity of EBT content and negatively associated with non-EBT content.

Supervision Session-Specific Factors
Intensity of supervision content is likely predicted by supervision session factors, specifically the overall time allocated to the supervision session and time allocated to any one client or case. Client caseloads in public mental health can be high. In the statewide initiative from which our sample was drawn, the average caseload was nearly 40 (19). Caseload size can limit EBT supervision time overall or time dedicated to any one case, which may in turn limit the possible intensity of supervision coverage of any single content area. In the objective coding study on which the current investigation builds (25), discussion of an EBT for any individual case averaged just under 12 min. We hypothesized that more time spent in supervision and more time per case would predict intensity of coverage for all three content elements (two EBT and one non-EBT).

METHODS
Data for the current study comes from a larger National Institute of Mental Health-funded study of workplace-based clinical supervision [see study protocol: (50)]. Participants were part of a state-funded EBT training initiative in public mental health in Washington State, which provides yearly in-person training and 6 months of expert consultation for TF-CBT [for more details on the training approach see (51)]. The current study uses objectively coded audio recordings of supervision collected during the “supervision as usual,” descriptive phase of the larger study and from baseline self-report surveys, prior to a subsequent randomized controlled trial (RCT) of two supervision approaches.

Procedure
The overall procedure was that supervisors and clinicians provided consent and completed a measures battery at baseline. Over the course of the following year, supervisors audio recorded all of their supervision sessions and these were coded.

The study team first identified organizations that had participated in the state-funded EBT initiative and had at least one TF-CBT-trained supervisor still at the organization. Supervisors who agreed to participate then identified eligible clinicians from among their supervisees. The study team contacted these clinicians to invite their participation and obtain informed consent. Of those approached, 72% of the organizations, 76.7% of the supervisors, and 76% of the clinicians consented to participate.

Data Collection
Supervisor and clinician study participants completed one online self-report survey at the beginning of the study before participating in a 2-day TF-CBT booster and study procedures training. Both clinicians and supervisors received $30 for completing the surveys. Supervisors who participated in the study were asked to audio record the portions of their individual supervision sessions that pertained to participating clinicians’ TF-CBT cases for one year (October, 2012–September, 2013). All audio recordings of these supervision sessions were sent to the study team. Supervisors did not record informal supervision sessions that occurred outside of regular supervision time or group supervision sessions. The audio recordings were saved on study-provided, password-protected tablet devices. The recordings were transferred to the study team using a cloud-based server compliant with the Health Insurance Portability and Accountability Act of 1996. Organizations that participated received $3,000 at the end of the RCT study.

The Washington State Institutional Review Board approved all study procedures.

Participants
Supervisors
Table 2 presents demographic information for all participants. Participants for these analyses included 28 supervisors who submitted audio recordings, representing 17 public mental health organizations located in 23 separate offices. In order to meet study inclusion criteria, participants were required to have received TF-CBT-specific training as part of the EBT initiative, to be a current supervisor of a clinician in the study, to be currently employed at a public mental health organization, and to have no immediate plans to leave the organization. An additional 5 supervisors participated in this phase of the study but did not submit audio recordings and were therefore excluded from these analyses. As described elsewhere (25), there were few significant differences between supervisors who submitted or did not submit recordings, except that supervisors who submitted recordings were slightly older, more likely to endorse CBT as their primary theoretical orientation, and less likely to endorse family systems therapy or art/play therapy.

Clinicians
Participants included 70 clinicians who were recorded in supervision sessions. Eligibility criteria for clinicians to participate in the study included: trained in TF-CBT through the statewide initiative, currently provide TF-CBT to children and adolescents, supervised by a supervisor involved in the study, employed at least 80% full-time equivalent or more, no immediate plans to leave the organization, and provided therapy in English (to enable coding of TF-CBT fidelity for other analyses). An additional 15 clinicians participated in this phase
Supervisors and clinicians completed the six-item Evidence-Based Organizational Checklist to assess the level to which their organizations expect, support, and reward EBT. All participant scores within each organization were aggregated to create an organizational implementation climate score. The content in this measure is similar to that of another implementation climate measure that was not available when the study began (52). Items are rated on a 4-point Likert scale (1, never; 2, occasionally; 3, most of the time; 4, ongoing/ routine). Example items from this measure include, “Executive leadership (e.g., administrators, directors) explicitly and repeatedly express support for and promote use of EBT,” and “Clinicians are provided with EBT training opportunities and ready access to EBT materials (manuals, handouts, equipment).” Previous studies have verified the unidimensionality and internal reliability of measure scores [see (51)]; the current study replicated good internal reliability (Cronbach’s α = 0.86). Higher scores indicate a more supportive EBT implementation climate. Construct validity of the measure is supported by a significantly high office-level Intraclass Correlation ICC(1,1) of 0.41. We use the ICC here to indicate “validity” rather than “reliability” because the clustering of implementation climate ratings by members of the same office indicates that climate is a shared perception at the office level (53, 54). Due to the small number of supervisors per office and challenges with a four-level model, we included implementation climate in analyses at the supervisor level (e.g., two supervisors in the same office would have the same climate score).

Participant Characteristics
Supervisors and clinicians were asked to provide information on their age, sex, ethnicity, race, number of years they had conducted therapy, and whether they felt they mainly used EBTs in their work. Participants indicated the total number of different types of training experiences they had with TF-CBT out of 12 possible options (e.g., “completed a 2-day in-person training,” “read the 2006 TF-CBT book”); experiences were summed. Participants endorsed their primary theoretical orientation from a list of 2006 TF-CBT book); experiences were summed. Participants endorsed their primary theoretical orientation from a list of 10 possible options. Supervisors provided an estimate of the percentage of time they spent providing supervision, whether they still actively performed clinical work, and chose the TF-CBT element that they felt was most difficult to supervise, which we transformed into a variable indicating whether or not they chose exposure as the most difficult element to supervise.

TF-CBT Self-Efficacy
Supervisor and clinician self-efficacy in TF-CBT was assessed using an 11-item index adapted from two previous measures (55, 56). Participants rated their level of competence implementing TF-CBT on a 5-point Likert scale (0, not at all; 1, a little bit; 2, somewhat; 3, very much; 4, exceptionally) using items such as “Completing trauma narratives with children,” and “Analyzing complex clinical situations from a TF-CBT perspective.” An exploratory factor analysis using maximum likelihood extraction

of the study, but audio recordings of their supervision sessions were not submitted, and they were therefore excluded from the study. As reported elsewhere (25), clinicians who were recorded and not recorded differed only on a few variables: clinicians who were recorded had provided psychotherapy for longer and were less likely to have a degree in Marriage and Family Therapy.

Measures
Below, we describe the measures used in this study. For additional information, see the measures table in Appendix A in Supplementary Material.
in the current sample justified retaining a single factor accounting for 56% of the variance; Cronbach's alpha was 0.92.

**Declarative Knowledge and Skill With TF-CBT and Exposure**

The Skill in Implementing Components: Trauma and PTSD scale was used with supervisors and clinicians to obtain the self-reported understanding of and skill in the major components of CBT for cases with trauma and PTSD (51). It includes 11 items rated on a 6-point scale ranging from 0 (do not use) to 5 (advanced), and asks participants to rate their understanding and skill of elements such as “Psychoeducation” and “Cognitive Coping.” In psychometric testing using an earlier version of this measure that asked about other elements in addition to trauma and PTSD, the trauma and PTSD scale emerged as a clear factor (51). Data from the current study had very high internal consistency (Cronbach's α = 0.91). We used the total mean score as well as a mean of two items, “in vivo Exposure” and “Trauma narrative.”

**TF-CBT Knowledge**

Supervisors and clinicians completed a 13-item multiple choice test of TF-CBT knowledge that combines items from the Denver Post Health Survey (57) with items added by our team, and includes content similar to the knowledge test used for the clinician TF-CBT certification program (https://tfcbt.org). Participants provided multiple choice or true/false response ratings to items such as “When teaching cognitive coping, wait to challenge distorted/unhelpful cognitions related to trauma.” The measure has been found to have a good response range for item difficulty and item discrimination, and has demonstrated convergent validity with number of trainings and TF-CBT self-efficacy (19).

**EBT Attitudes**

Supervisors and clinicians completed the Modified Practice Attitudes Scale (MPAS) to assess attitudes toward EBTs (58). The current study used a five-item version of the MPAS with acceptable internal consistency and good validity (59). Participants indicated their agreement with statements such as “Clinical experience and judgment are more important than using evidence-based treatments,” using a 4-point scale ranging from 0 (not at all) to 4 (to a very great extent). The current study found acceptable internal consistency (Cronbach's α = 0.78).

**Supervision Session Time**

During coding of audio recordings (described below), coders determined the length of the supervision session (in minutes) and number of cases discussed. Average minutes per case was calculated by dividing the total session time by the number of cases.

**Supervision Content**

The dependent variables used in this study, i.e., intensity of supervision content areas, were obtained using the Supervision Process Observational Coding System (SPOCS), which was adapted from the Therapeutic Process Observational Coding System for Child Psychotherapy—Strategies scale [TPOCS-S; (27, 60)]. The TPOCS-S categorizes psychotherapy treatment intervention elements using direct observation. Similarly, the SPOCS categorizes supervision elements, applying Garland et al.’s (5) adaptation of the TPOCS-S by stratifying codes into content and technique domains. For the current study, we focused only on the content domain.

There are 16 content areas in the SPOCS, described in detail elsewhere (25) and in Appendix A in Supplementary Material. We examined three content items for the purposes of the current paper: exposure [which combines two exposure elements: (a) trauma narration and processing and (b) in vivo mastery of trauma reminders, see Table 1], assessment, and other topics (including crisis or case management). As reported in Dorsey et al. (25) trained coders rated content in 5-min intervals and then considered ratings across intervals to generate an overall intensity score for each individual content item. These three content items had normally distributed intensity scores, with ratings from 0–6, (0 = not present, 1–2 = low intensity, 3–4 medium, 5–6 = high). For instance, low intensity ratings for assessment reflected brief mentions of the content (e.g., “Don’t forget to do the weekly assessment”). High intensity ratings for assessment reflected more in-depth discussion, such as planning for assessment in an upcoming session (rationale, strategies to remove barriers), in treatment generally, and/or review of assessment results (e.g., scores, clinical significance, change over time) and implications for the treatment plan (such as whether assessment scores indicate that a specific component is warranted). As the SPOCS is newly developed and there is no existing measure with which to compare, we lack complete psychometrics on this measure. However, as described below, the coding team achieved very high Interrater reliability, which suggests that the SPOCS identifies distinct and observable components.

**Coder Training and Session Sampling**

**Coder Training**

The details of SPOCS coder training are described elsewhere (25). Coders were six post-baccalaureate research assistants. Coders attended an initial training, which included content review, group coding, and detailed coding manual review and discussion. They then coded 10 training files independently to ensure satisfactory Interrater reliability across group members and with the last author. Official coding for the study began once each coder's ratings reached an established criterion: Interrater reliability using two-way random single measure intraclass correlation coefficients [ICC(2,1) ≥ 0.80; (61)]. For individual content/technique items for which an ICC(2,1) ≤ 0.60, coders were required to engage in additional practice and review. Coders were required to re-read the coding manual monthly, discuss, and reference the manual when questions or confusion arose, and attend recurring booster trainings to prevent drift. Coders were randomly assigned supervision files. Possible rater drift was monitored through masked coders double-coding sessions at regular intervals; ICCs remained strong throughout and no coder fell below an ICC(2,1) of 0.80.
Session Sampling Procedures
In total, we received 638 supervision recordings. Per supervisor, up to 23 individual supervision sessions were coded (when available), resulting in 438 coded recordings. When a supervisor submitted 23 or fewer files, the study team coded all submitted files. When a supervisor submitted over 23 files, 23 files were randomly selected using a form of stratified random sampling in which selected recordings were distributed across time and participating clinicians.

Interrater Reliability
To test interrater reliability, 105 (23.9%) of the 438 sampled session recordings were coded by multiple coders. The overall group average ICC assessing reliability was ICC(2,6) = 0.87, representing excellent reliability (61). Coders had excellent individual ICCs of 0.84 or higher. At the item level, ICC(2,1) statistics ranged from good to excellent, Exposure = 0.92, Assessment = 0.76, Other topics = 0.85.

Analyses
Analyses were conducted in SPSS 19. Means, standard deviations, and percentages were calculated for participant descriptive information and content items. Using null models with no predictors, three separate 3-level mixed effects models with random intercepts at the supervisor and clinician level (supervision session nested within clinician nested within supervisor) were used to compute intraclass correlations (ICCs), which are the proportion of variance for each dependent variable attributable to each level. Although 4-level models that include nesting within organization would be more appropriate, several organizations had only a single supervisor participating in the study, and therefore, clustering estimates for these models failed to converge. Restricted Maximum Likelihood (REML) estimation and an unstructured covariance matrix were used to obtain final parameter estimates.

Model building for hypothesis testing followed standard protocol (62). For each unique independent-dependent variable combination, a separate model was computed, which tested the unique bivariate relationship between each independent and dependent variable, similar to the standard practice of computing a correlation table prior to ordinary least squares regression modeling. Based on these analyses, we built models beginning with level 1. All predictor variables were entered as grand mean centered to aid interpretation of the intercept—using this approach, the intercept represents the estimated mean score of the dependent variable, rather than the estimated score if all predictors were zero. Level-1 and level-2 predictors were entered in bivariate analyses as fixed effects and then as random effects. In all models below, no randomly varying slopes were significant, and allowing the effects of these level-1 predictors to vary did not improve model fit, or models failed to converge; thus, all level-1 and level-2 slopes in all models were fixed. We removed or retained parameters based on model fit statistics, assessed using significance of −2 log likelihood deviance and magnitude of Bayesian Information Criteria (BIC) deviance, with values for the BIC above 2 considered positive evidence of model superiority, and values above 10 indicating strong evidence (63). After a level-1 model was built, each level-2 predictor that had been significant at \( p < 0.05 \) during the bivariate testing described above was added as a fixed effect in a stepwise fashion to assess model fit. When two or more variables were individually significant but non-significant when jointly entered, model fit statistics were used to determine the best fitting parsimonious model.

RESULTS

Exposure Contextual Information; Dorsey et al. (25) Analyses
As originally reported in the study on which our investigation builds (25), exposure was frequently covered, in that it was mentioned in 82% of the coded supervision sessions. The intensity of exposure coverage varied, however. In 17% of sessions it was not discussed, in 24% of the sessions it was discussed with low intensity, in 41% with medium intensity, and in 17% with high intensity (\( M \) intensity across sessions = 2.64, \( SD = 1.75 \)). A null (no predictor) model predicting intensity of exposure coverage indicated that 16% of the variance in exposure coverage was at the supervisor level and 19% at the clinician level, with the remaining 65% at the individual supervision-session level. Therefore, intensity of supervision time spent on exposure appeared to be attributable to factors at both the supervisor and clinician levels.

Current Analyses
Item range, means, standard deviations for each predictor variable are depicted in Table 3. Bivariate models for each potential predictor of intensity of exposure coverage at the organization, supervisor, clinician, and supervision-session level resulted in few significant associations (see Table 3). Longer TF-CBT supervision sessions, more supervision time per case, supervising a clinician with a cognitive behavioral theoretical orientation, and a more positive organizational implementation climate were associated with greater intensity coverage of exposure in supervision. Supervisors’ belief that exposure or the trauma narrative was the most difficult element to supervise was associated with lower intensity of exposure coverage.

The final model predicting exposure is depicted in Table 4. Average estimated exposure intensity was 2.7. Time spent per case was significantly and positively associated with exposure intensity, with each additional minute of time related to a 0.04 increase in intensity. Implementation climate was also significantly and positively associated, with each additional one-point increase in implementation climate associated with a 0.87 increase in exposure intensity. Therapists with a CBT orientation had an average exposure intensity score 0.52 points higher than those with another orientation, as indicated by improved model fit (\( \Delta -2LL_{(1)} = 37.6, p < 0.001 \)) and strong evidence of model superiority (\( \Delta BIC_{(1)} = 37.7 \)) when compared to a model without this variable. However, the individual variable parameter did not meet statistical significance (\( p = 0.112 \)), so cautious interpretation is warranted. The final model accounted for 12.5% of the overall variance. This included 3.5% of
### TABLE 3 | Predictor descriptives and mixed linear model coefficients showing bivariate associations among supervision content and characteristics of the supervisor, clinician, and supervision session.

<table>
<thead>
<tr>
<th>ORGANIZATIONAL CHARACTERISTICS</th>
<th>Range</th>
<th>Mean (SD)</th>
<th>Exposure</th>
<th>Assessment</th>
<th>Other topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation climate</td>
<td>1.8–3.8</td>
<td>3.1 (0.53)</td>
<td>1.039*</td>
<td>0.818*</td>
<td>−0.654</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SUPERVISOR CHARACTERISTICS</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

#### Expertise
- # of TF-CBT trainings: 1–10, 4.7 (2.0), 0.124, 0.153, −0.260*
- Primarily uses EBT: 0–1, 0.75 (0.44), 0.677, 0.797*, −0.639
- TF-CBT knowledge test: 6–13, 10.1 (1.7), 0.140, 0.100, −0.302*

#### Beliefs in own ability
- TF-CBT efficacy: 2.5–4.6, 3.6 (0.49), 0.804, 0.446, −0.273
- Declarative TF-CBT knowledge/skill: 2–5, 3.8 (0.69), −0.038, 0.246, −0.204
- Declarative exposure knowledge/skill: 2–5, 3.6 (0.93), 0.061, NA, NA

#### Willingness
- EBT attitudes: 0–1, 0.52 (0.50), −0.505*, NA, NA

#### Theoretical orientation
- Cognitive behavioral: 0–1, 0.68 (0.47), 0.015, 0.933*, −0.189
- Family systems: 0–1, 0.18 (0.39), −0.256, −0.859*, −0.084

#### Other practice characteristics
- Number of years conducting therapy: 3–37, 12.1 (6.8), 0.025, −0.010, −0.042
- % time providing supervision: 5–90, 37.5 (20.1), 0.005, 0.001, 0.006
- Currently provides clinical services: 0–1, 0.86 (0.35), 0.570, 0.533, 0.570

<table>
<thead>
<tr>
<th>CLINICIAN CHARACTERISTICS</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

#### Expertise
- Number of TF-CBT trainings: 1–9, 3.3 (1.76), −0.013, −0.044, 0.067
- Number of years conducting therapy: 1–30, 4.8 (5.01), −0.007, 0.007, −0.006
- Primarily uses EBT: 0–1, 0.77 (0.42), 0.391, 0.120, 0.264
- TF-CBT knowledge test: 3–13, 9.0 (1.90), 0.021, −0.071, −0.044

#### Beliefs in own ability
- TF-CBT efficacy: 1–5, 3.1 (0.70), 0.051, −0.076, 0.062
- Declarative TF-CBT knowledge/skill: 1–5, 3.1 (0.80), 0.043, 0.056, 0.196
- Declarative exposure knowledge/skill: 1–5, 3.0 (1.05), 0.087, NA, NA

#### Willingness
- EBT attitudes: 2.4–5, 3.9 (0.49), 0.179, 0.243, 0.086

#### Theoretical orientation
- Art therapy: 0–1, 0.01 (0.11), −0.395, −0.140, −0.676
- Cognitive behavioral: 0–1, 0.65 (0.48), 0.745*, 0.388, −0.312
- Family systems: 0–1, 0.09 (0.29), −0.272, 0.562, −0.049
- Humanistic: 0–1, 0.06 (0.23), 0.435, −0.676, 0.080
- Play therapy: 0–1, 0.03 (0.18), −0.458, −0.233, 0.017
- Psychodynamic: 0–1, 0.06 (0.27), −0.904, −0.313, 0.716

<table>
<thead>
<tr>
<th>SUPERVISION SESSION CHARACTERISTICS</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

#### Supervision session duration in minutes
- 1–72, 21.6 (15.0), 0.025*, 0.009, 0.051*
- Minutes per case
  - 1–51, 12.2 (8.6), 0.034*, 0.006, 0.025*

* $p < 0.05$

*Each supervision content area refers to the intensity with which the clinical content was discussed during supervision sessions. Exposure is defined as discussions of a technique to gradually reduce fears and anxiety bysubjecting the client to a feared stimulus, such as memories of a traumatic event. Assessment is defined as discussions of information about the child’s psychiatric symptoms or behavior problems from standardized, formal assessment measures and functional analysis. Other topics is defined as discussions of issues unrelated to the child’s traumatic experiences or not directly related to TF-CBT components.*

the variance at the individual supervision-session level and 66.5% of the variance at the supervisor level. Variance at the clinician level slightly increased from the null to the final model.

**Assessment**

As reported elsewhere (25), compared to exposure, assessment was more rarely discussed, and included in only 55% of the coded supervision sessions. The intensity of assessment coverage varied,
TABLE 4 | Mixed linear model predicting intensity of exposure coverage in workplace-based supervision.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervision session minutes per case</td>
<td>0.04</td>
<td>0.01</td>
<td>3.99</td>
</tr>
<tr>
<td>Therapist: CBT orientation</td>
<td>0.52</td>
<td>0.32</td>
<td>1.61</td>
</tr>
<tr>
<td>Implementation climate</td>
<td>0.87</td>
<td>0.32</td>
<td>2.73</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variance components</th>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>p</th>
<th>% variance accounted for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual (session duration)</td>
<td>1.97</td>
<td>0.15</td>
<td>13.41</td>
<td>&lt; 0.001</td>
<td>3.5%</td>
</tr>
<tr>
<td>Supervisor</td>
<td>0.17</td>
<td>0.22</td>
<td>0.75</td>
<td>0.456</td>
<td>66.5%</td>
</tr>
<tr>
<td>Clinician</td>
<td>0.60</td>
<td>0.25</td>
<td>2.41</td>
<td>0.016</td>
<td>-1.7%</td>
</tr>
<tr>
<td>Overall</td>
<td>2.73</td>
<td>0.15</td>
<td>18.15</td>
<td>&lt; 0.001</td>
<td>12.5%</td>
</tr>
</tbody>
</table>

with only a few sessions addressing it with high intensity (in 45% of the sessions it was not discussed, in 32% it was discussed with low intensity, in 18% with medium intensity, and in 5% with high intensity; M intensity across sessions = 1.30, SD = 1.52). A null model indicated that 23% of the variance clustered at the supervisor level, only 2% clustered at the clinician level, and the remaining 75% of the variance was at the individual supervision-session level, implying that clinician-level factors likely do not account for any significant amount of assessment coverage during supervision.

Current Analyses

Consistent with this, bivariate models found that assessment was significantly associated only with supervisor-level variables. Higher assessment intensity scores were associated with supervisors who reported that they primarily used EBTs, more positive implementation climate, and supervisors who reported having a CBT orientation. Lower assessment intensity was associated with supervisors who reported that they primarily used EBTs, more positive implementation climate, and supervisors who reported having a family systems theoretical orientation.

The final model for assessment is depicted in Table 5. Average estimated assessment intensity was 1.3, and it was predicted only by implementation climate (each one-point increase in implementation climate was related to a 0.64 increase of intensity). The model accounted for 7.3% of the overall variance, mostly due to accounting for 32.8% of the variance specifically at the supervisor level; clinician variance slightly increased.

TABLE 5 | Mixed linear model predicting intensity of assessment coverage of workplace-based supervision.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation climate</td>
<td>0.82</td>
<td>0.26</td>
<td>3.16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variance components</th>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>p</th>
<th>% variance accounted for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual (session duration)</td>
<td>1.76</td>
<td>0.13</td>
<td>13.64</td>
<td>&lt; 0.001</td>
<td>0%</td>
</tr>
<tr>
<td>Supervisor</td>
<td>0.36</td>
<td>0.13</td>
<td>2.38</td>
<td>0.017</td>
<td>32.8%</td>
</tr>
<tr>
<td>Clinician</td>
<td>0.05</td>
<td>0.08</td>
<td>0.56</td>
<td>0.575</td>
<td>-6.0%</td>
</tr>
<tr>
<td>Overall</td>
<td>2.16</td>
<td>0.15</td>
<td>13.41</td>
<td>&lt; 0.001</td>
<td>7.3%</td>
</tr>
</tbody>
</table>

TABLE 6 | Mixed linear model predicting intensity of “other topics” coverage in workplace-based supervision.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor Session Duration</td>
<td>0.05</td>
<td>0.01</td>
<td>11.22</td>
</tr>
<tr>
<td>Supervisor TF-CBT Knowledge</td>
<td>-0.17</td>
<td>0.06</td>
<td>-2.72</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variance components</th>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>p</th>
<th>% variance accounted for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual (session duration)</td>
<td>1.09</td>
<td>0.08</td>
<td>13.57</td>
<td>&lt; 0.001</td>
<td>14.0%</td>
</tr>
<tr>
<td>Supervisor</td>
<td>0.05</td>
<td>0.08</td>
<td>0.72</td>
<td>0.473</td>
<td>92.6%</td>
</tr>
<tr>
<td>Clinician</td>
<td>0.18</td>
<td>0.09</td>
<td>20.59</td>
<td>0.039</td>
<td>-1.0%</td>
</tr>
<tr>
<td>Overall</td>
<td>1.32</td>
<td>0.15</td>
<td>13.41</td>
<td>&lt; 0.001</td>
<td>39.2%</td>
</tr>
</tbody>
</table>

with medium intensity, 27% discussed with high intensity). A null model found that 34% of the variance in intensity of other topics was at the supervisor level, 8% was at the clinician level, and the remaining 58% was at the individual supervision session level.

The “other topics” that were discussed consisted of case background information (45%), administrative work (15%), case management (10%), child symptoms and behavior problems (10%), non-trauma focused treatment elements (10%), non-work related conversations (8%), and crisis management (2%). Bivariate models found that intensity of supervisory time spent on other topics was predicted by duration of the session, minutes per case, lower supervisor scores on the TF-CBT knowledge test, and less supervisor-reported training in TF-CBT.

Current Analyses

The final model for other topics is depicted in Table 6, and estimated the average intensity of supervisory time spent on other topics at 3.5. Other topics was predicted by the duration of the supervision session (each additional minute was associated with a 0.05 increase of other topic intensity), and supervisors with

Other Topics

Contextual Information; Dorsey et al. (25) Analyses

As reported elsewhere (25), other topics was the content item delivered in almost every coded supervision session (96%). It was covered with the greatest mean level intensity (3.46, SD = 1.47), although intensity did vary across coded sessions (4% not discussed, 19% discussed with low intensity, 50% discussed with medium intensity, 27% discussed with high intensity). A null model found that 34% of the variance in intensity of other topics was at the supervisor level, 8% was at the clinician level, and the remaining 58% was at the individual supervision session level.

The “other topics” that were discussed consisted of case background information (45%), administrative work (15%), case management (10%), child symptoms and behavior problems (10%), non-trauma focused treatment elements (10%), non-work related conversations (8%), and crisis management (2%). Bivariate models found that intensity of supervisory time spent on other topics was predicted by duration of the session, minutes per case, lower supervisor scores on the TF-CBT knowledge test, and less supervisor-reported training in TF-CBT.

Current Analyses

The final model for other topics is depicted in Table 6, and estimated the average intensity of supervisory time spent on other topics at 3.5. Other topics was predicted by the duration of the supervision session (each additional minute was associated with a 0.05 increase of other topic intensity), and supervisors with
lower scores on the TF-CBT knowledge test spent more time on other topics (each additional point on the knowledge test was associated with a 0.17 decrease in other topic intensity). Overall, the model accounted for 39.2% of the variance in other topics intensity. This included 14% of the variance at the supervision session level and 92.6% of the variance at the supervisor level.

**DISCUSSION**

Workplace-based supervision might be an effective dissemination and implementation strategy to increase the adoption, fidelity, and sustainment of EBTs. This study used an innovative method of coding supervision elements to explore the predictors of EBT content delivery during workplace-based supervision. To our knowledge, this is the first study to use objectively coded data from workplace-based supervision of EBT to explore predictors of intensity of coverage of EBT content. We found some support for multi-level predictors at the organization and supervisor levels, and as hypothesized, time played an important role. However, surprisingly few supervisor-level and no clinician-level variables predicted intensity of coverage for either of the two EBT content areas (i.e., exposure and assessment) or for other topics in multivariate models, and a large amount of variance remained unexplained for the two EBT content areas.

Implementation climate predicted intensity of coverage for both exposure and assessment. Overall, results suggest that a climate that supports, expects, and rewards EBT use may be one of the most important factors for improving the degree to which supervisors cover EBT in their supervision sessions. This finding is in line with two other supervision content-focused studies in which implementation climate was a significant predictor of clinician-reported supervision time spent on case conceptualization and interventions (19) and intensity of TF-CBT content (23). The present study is a constructive replication (64) of this prior work in that it analyzes objectively coded data instead of self-report, providing stability for these findings. Of the variables we explored, implementation climate was the strongest predictor. After controlling for other significant covariates, each one-point increase in climate was associated with a nearly one-point increase in intensity of exposure and assessment coverage, on a 6-point scale. These findings indicate that supervisors in organizations with more positive implementation climates may be more likely to provide the fidelity monitoring and support necessary as an inner context support during the latter two EPIS phases, active implementation and sustainment (10). The findings highlight the importance of creating an environment within which supervisors feel supported to carve out supervision time to cover EBT in greater intensity and feel that this is expected and rewarded in their organizations, despite competing demands on limited supervision time in the context of clinicians’ high caseloads.

In light of our findings, it is important to consider that the association between supervision and implementation climate is likely bidirectional; supervisors both create and are shaped by implementation climate. Other studies on clinical supervision have raised similar questions surrounding this bidirectionality. For example, an observational study which adapted a supervision model from Multisystemic Therapy to implement social-emotional interventions in schools (65) raised questions about “the extent to which the scope of clinical supervision, and responsibility of the clinical supervisor, extends to the proactive cultivation and maintenance of organization-intervention fit…” (p. 55). Relatedly, Birken and colleagues proposed that “middle managers,” defined as employees who supervise frontline staff and are themselves supervised by top organizational leaders, play several key roles hypothesized to positively impact implementation climate and implementation effectiveness (66). As middle managers, supervisors go beyond providing clinical oversight, and regularly support EBT implementation at their organizations. Studies testing the impact of supervisor-level interventions on implementation climate and effectiveness are currently underway [e.g., (67)], and more studies are needed to further unpack the relationship between supervision and implementation climate.

Our hypotheses that time would be positively related to coverage intensity were supported for exposure and other topics but, interestingly, not for assessment. Among other determinants of practice, the lack of time is frequently endorsed by clinicians and other healthcare providers as a substantial barrier to EBT implementation [e.g., (68–71)]. The role that time plays appears to be complex. For exposure, more time per case was a stronger predictor than time allotted to the EBT supervision overall. In the objective coding study on which the current study builds, the average supervision time dedicated to a specific case was just over 12 min (25). Assuming a linear relationship, supervision time for any individual case would need to be doubled, from 12 to 24 min, to obtain a 1-point increase in intensity of exposure coverage in supervision. For the intensity of coverage of other topics, time allotted to the EBT supervision overall was a stronger predictor than time per case. For every minute increase in session duration, we saw a small (0.05) increase in the intensity of other topics. Most of the content areas that comprise other topics were related to the case being supervised, and primarily included discussion of case background information (about half of the content). However, off-topic or administrative content made up nearly a quarter of the other topics’ content. Of all 16 areas coded, the variance attributable at the supervisor-level was the highest for other topics [34%; (25)]; therefore, more time may permit supervisors to focus on other topics beyond the EBT with greater intensity. Questions remain regarding this relationship: it could be that other topics conversations in supervision lead to lengthier supervision sessions, or it could be that shorter supervision time has the effect of enabling greater efficiency and strategic use of time to focus more on EBT.

Only one supervisor characteristic was significantly associated with any of our dependent variables after controlling for other variables. Supervisors with less knowledge of the specific EBT (TF-CBT) covered other topics with greater intensity. This finding makes logical sense, as EBT expertise would seem to be a requirement for greater intensity of coverage. Other significant bivariate associations, although not supported in the multivariate models, also suggest that supervisor-specific
expertise and experience may play a role in intensity of content coverage (e.g., lower supervisor comfort supervising exposure was associated with less intense exposure coverage; supervisors primarily using EBTs and having a CBT theoretical orientation were associated with greater intensity of assessment coverage; supervisors with more TF-CBT training were associated with less intense other topics coverage). However, our hypotheses about multiple other variables being associated with content had no support even at the bivariate level (e.g., TF-CBT-specific self-efficacy, declarative knowledge and skill with TF-CBT, years of experience).

Similarly, the lack of empirical support for clinician-level predictors for any of the three content elements was unexpected. Neither of the variables capturing clinician expertise with the EBT was significantly associated with coverage intensity for any of the three supervision content elements. This was particularly surprising for exposure, which was unique among the three content areas in that it clustered with more than 10% of the variance at both the clinician and supervisor levels. Even in bivariate analyses, there was only one significant clinician-level predictor, and it was associated with willingness to address EBT: clinicians' self-report of a CBT theoretical orientation was related to more intense exposure coverage. This suggests that while the makeup of supervision may be driven in some small part by contributions of the clinician and the supervisor, the tailoring of supervision content to the needs of the clinician (e.g., based on skill or experience) may occur less frequently than our team predicted.

While our model for other topics was strongly predictive, with supervisor knowledge and supervision session length explaining nearly 40% of the overall variance, models for exposure and assessment did not explain much variability (13 and 7%, respectively). Based on these ICCs, it appears that much of the variance in delivery of these two EBT content elements occurs at the supervision-session level. There are three possible reasons for variability at this level: general random error, measurement error associated with coding reliability, and true session-level differences; unfortunately, these sources of variability are not statistically separable. The very high coder interrater reliability indicates that measurement error is not likely to be a major source of variance. Therefore, session-level sources of variability likely arise from multiple variables for which we do not have data, including the specific session-level needs of the clients, the timeline of treatment (e.g., assessment may be more likely to be discussed early in the treatment process), and the moods and cognitive loads of clinicians and supervisors during any specific supervision session. These variables can act as statistical noise, creating challenges for detecting predictors.

This study has a number of strengths. Findings are backed by strong internal validity from the use of our objective coding measure for supervision, and they replicate findings from analyses using self-reported data. Supervision session data was obtained from actual workplace-based supervision sessions from participants in a statewide EBT implementation initiative, providing generalizability to other EBT implementation efforts that attempt to leverage workplace-based supervision of EBTs. However, there were some limitations. Many of the content elements we coded (e.g., coverage of cognitive processing, clinician modeling, and role-play during treatment sessions) were not analyzed due to limited variability (i.e., rare occurrence; occurred with low intensity). Although the sample size for number of recordings was high (438 recordings), due to the nested nature of the data, the sample size for supervisors and clinicians (28 and 70, respectively) limited our power to detect effects at these levels. Also, as described previously, many variables that explain session-level variance were unmeasured (e.g., client needs/progress, clinician/supervisor temporal mood). The data are correlational and causal direction cannot be demonstrated. Additionally, to protect sensitive client information, supervisors were asked to only record the portion of individual supervision that pertained to TF-CBT cases. Similarly, we did not sample or code informal “drop-in” or group-based supervision, all of which may also contribute to the supervision of any one case.

Considering practical implications for workplace-based clinical supervision as a support for EBT implementation, our findings suggest that spending more time on supervision may not be the most efficient method to heighten the EBT focus of supervision. Time was not a significant predictor for assessment, and to increase intensity of coverage for exposure, a substantial and likely infeasible amount of time would need to be added. Simply increasing the amount of time for supervision might also result in supervision that is less focused on EBT content (i.e., more time focused on other topics). In contrast, having a more positive implementation climate had a strong effect. This suggests that efforts to improve the degree to which individuals perceive that their organization supports, expects, and rewards EBT use may positively impact the EBT focus of supervision and in turn support higher fidelity EBT delivery by clinicians. However, the field is only beginning to examine practical and effective methods for enhancing implementation climate. Future research on impacting supervision structure could explore the feasibility of improving implementation climate and the nature and direction of the relationship between organizational climate and supervisor behaviors. Meanwhile, additional research could be conducted to better identify the supervisor, clinician, and client-level variables that explain EBT and other content coverage in supervision.

**ETHICS STATEMENT**

This study was carried out in accordance with the recommendations of the Washington State Institutional Review Board, within the State of Washington Department of Social and Health Services. The protocol was approved by the Washington State Institutional Review Board. All participants gave written informed consent in accordance with the Declaration of Helsinki.

**AUTHOR CONTRIBUTIONS**

MP was responsible for conceptualizing, analyzing, and writing this manuscript. LL, JH, PM, and KB were responsible for conducting project activities and writing the manuscript. LL and KB also coded supervision audio files. SD and ED were
responsible for developing the supervision coding manual, directing project activities, conceptualizing, and writing.

FUNDING
Funding for this research project was supported by the National Institute of Mental Health (R01 MH095749, Dorsey, PI) and (F31MH109245, Harrison, PI).

ACKNOWLEDGMENTS
The authors would like to acknowledge the Washington State Division of Behavioral Health and Recovery for funding and supporting the Washington State TF-CBT and CBT+ Initiative and for being supportive of this research partnership. We thank all participating organizations, supervisors, and clinicians; the STEPS Team for facilitating data collection; and Bryce McLeod and Ann Garland for assistance in adapting the TPOCS-S for supervision sessions.

SUPPLEMENTAL MATERIAL
The Supplemental Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2018.00280/full#supplementary-material

REFERENCES


**Conflict of Interest Statement:** SD and ED have received honorariums for providing TF-CBT training.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pullmann, Lucid, Harrison, Martin, Deblinger, Benjamin and Dorsey. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Evidence-based practice (EBP) implementation requires substantial resources in workforce training; yet, failure to achieve long-term sustainment can result in poor return on investment. There is limited research on EBP sustainment in mental health services long after implementation. This study examined therapists’ continued vs. discontinued practice delivery based on administrative claims for reimbursement for six EBPs [Cognitive Behavioral Interventions for Trauma in Schools (CBITS), Child–Parent Psychotherapy, Managing and Adapting Practices (MAP), Seeking Safety (SS), Trauma-Focused Cognitive Behavior Therapy (TF-CBT), and Positive Parenting Program] adopted in a system-driven implementation effort in public mental health services for children. Our goal was to identify agency and therapist factors associated with a sustained EBP delivery. Survival analysis (i.e., Kaplan–Meier survival functions, log-rank tests, and Cox regressions) was used to analyze 19 fiscal quarters (i.e., approximately 57 months) of claims data from the Prevention and Early Intervention Transformation within the Los Angeles County Department of Mental Health. These data comprised 2,322,389 claims made by 6,873 therapists across 88 agencies. Survival time was represented by the time elapsed from therapists’ first to final claims for each practice and for any of the six EBPs. Results indicate that therapists continued to deliver at least one EBP for a mean survival time of 21.73 months (median = 18.70). When compared to a survival curve of the five other EBPs, CBITS, SS, and TP demonstrated a higher risk of delivery discontinuation, whereas MAP and TF-CBT demonstrated a lower risk of delivery discontinuation. A multivariate Cox regression model revealed that agency (centralization and service setting) and therapist (demographics, discipline, and case-mix characteristics) characteristics were significantly associated with risk of delivery discontinuation.
for any of the six EBPs. This study illustrates a novel application of survival analysis to administrative claims data in system-driven implementation of multiple EBPs. Findings reveal variability in the long-term continuation of therapist-level delivery of EBPs and highlight the importance of both agency and workforce characteristics in the sustained delivery of EBPs. Findings direct the field to potential targets of sustainment interventions (e.g., strategic assignment of therapists to EBP training and strategic selection of EBPs by agencies).

**Keywords:** evidence-based practices, sustainment, survival analysis, administrative claims data, children's mental health services

## INTRODUCTION

In response to a national call for an increased delivery of evidence-based practices (EBPs) in routine-care settings to improve the quality of care (1–4), mental health systems have increasingly mandated or incentivized the implementation of EBPs. As of 2014, more than 20 states have implemented evidence-based mental health therapies or medication practices either directly or through contracts with other organizations (5–8).

Evidence-based practice implementation requires substantial investments to support the mental health workforce (9, 10). Such costs are incurred through clinicians’ time spent on attending trainings (i.e., lost revenue for the agency) and costs to facilitate the supervision and fidelity monitoring of newly trained staff, including payments to external consultants or trainers (11). For example, staff training and supervision account for 24 and 17.3%, respectively, of the total costs associated with the implementation of Trauma-Focused Cognitive Behavioral Therapy (TF-CBT) at 10 community mental health (CMH) agencies (11). Other major costs identified included non-billable provider time (11.8%) and implementation-related team meetings (11.0%). In mental health services, workforce costs may be especially high given the complexity of multicomponent psychosocial EBPs, many of which have intensive certification requirements (12). As noted by Proctor et al. (13), these costs are dependent on the complexity, strategy, and setting of an intervention. As such, an even greater investment is needed when multiple complex interventions are rolled out in tandem in a given service system (14).

While there is a growing literature on the factors influencing the initial implementation of EBPs, less is known about what transpires in the years following their adoption (15–18). Failure to achieve long-term sustainment of adopted EBPs results in poor return on investment due to the limited public health impact of initiatives (19–21). It is therefore important to examine patterns of EBP sustainment and to identify agency and therapist factors that are associated with sustainment vs. discontinuation of practice delivery over time and inform the tailoring of implementation strategies.

Lack of EBP sustainment and limited success in future EBP implementation efforts are yoked, and potential barriers to practice sustainment abound. In some service settings, new EBPs tend to be cyclically adopted and de-adopted, based on local, state, or federal requirements, practice “trends,” and other contextual influences. Without effective sustainment, these serial EBP adoption and de-adoption cycles have immediate costs in terms of finances, person hours, and later downstream consequences, such as implementation overload and learned helplessness (22), staff cynicism, and resistance to innovation (23). However, two persistent conditions in public mental health service systems are key drivers of sustainment failures. First, stakeholders at the agency, system, and EBP developer levels have reported staff turnover to be the greatest barrier to sustainment (24). Second, inadequate long-term funding of implementation initiatives is a common challenge to EBP sustainment. In a study by Bond et al. (25) of the sustainment of five EBPs implemented across eight states, barriers to sustainment varied by site: 94% of discontinuing sites identified financial reasons as a major barrier, followed by 47% of discontinuing sites citing workforce factors (e.g., the availability of certified practitioners). By contrast, studying sustainment in the context of long-term stable funding of implementation may help to reveal other provider-level characteristics that promote EBP sustainment. Furthermore, when considering system-level outcomes, it is unclear whether EBP sustainment is maintained at the system level even when providers turn over across agencies or organizational units.

Sustainment has been measured in a variety of ways; moreover, sustainment outcomes may depend on how and when sustainment is assessed (17, 26). Possible sustainment outcomes include reach/penetration (i.e., the extent to which a practice is integrated in a service setting as a proportion of population served or population of providers delivering care) or service volume (i.e., the extent to which a practice is delivered over time in a number of agencies, therapists, children, and units of service) (27). Another major index of sustainment is the extent to which trained therapists continue to deliver an EBP to clients in a system once they are trained to do so.

As previously noted, staff turnover at agencies has been identified as a barrier to sustainment, representing penetration-related losses in training investments at the organization level [e.g., Ref. (28, 29)]. However, it is plausible that workforce investments may continue to yield benefits at the system level to the extent that therapists move between organizational units within a larger system that share fiscal resources. For example, Beidas et al. (10) found that 55% of the staff who left a CMH agency within 1 year of follow-up remained in the public sector system, whereas 35% acquired new jobs in the private sector. These findings suggest that even though EBP-training investments may be lost at the agency level when a staff member leaves the agency, some proportion of...
this loss may be recaptured within a system when the provider is retained at another unit within the system. The extent to which this occurs has not been studied and offers a complementary sustain- 
tment outcome that is particularly relevant to system-driven implementa- 
tion efforts.

Administrative claims data have been identified as a valuable resource for researchers and policymakers alike to understand the sustainment of mental health policy/program initiatives across agencies (15, 30, 31, 32, 33). Claims data provide an opportunity for a novel application of survival analysis to examine sustainment in the context of large, system-driven EBP implementations. Also known as duration analysis, event history analysis, or failure analysis, among other names, survival analysis is an analytic method used in a variety of fields, ranging from economics to sociology to engineering to measure the length of time until a defined event occurs (34). In mental health research, survival analysis has been used to examine therapist turnover [e.g., Ref. (29)] and client attrition/psychotherapy termination at a clinic [e.g., (35)], but has yet to be harnessed to study EBP sustainment in mental health services. Administrative claims data also provide the opportunity to identify potential factors associated with a sustained EBP delivery. It is plausible, for example, that certain therapist characteristics (e.g., bilingual competence) and case-mix characteristics (e.g., alignment of predominant population diagnosis to EBP) may bode well for a sustained EBP delivery.

**MATERIALS AND METHODS**

**Context**

The current study is an exploratory one that applies survival analysis to a novel context—measuring therapists’ sustained delivery of multiple EBPs in the context of a system-driven implementation. Administrative data were collected through the Prevention and Early Intervention (PEI) program in the Los Angeles County Department of Mental Health (LACDMH), the largest county mental health department in the USA (27). The purpose of this study was to use EBP-specific claims data to (1) characterize therapists’ continued vs. discontinued delivery of six EBPs [Cognitive Behavioral Interventions for Trauma in Schools (CBITS), Child–Parent Psychotherapy (CPP), Managing and Adapting Practice (MAP), Seeking Safety (SS), TF-CBT, and Positive Parenting Program (Triple P)] and (2) identify factors associated with a sustained EBP delivery. Consistent with implementation models [e.g., Ref. (26)], organizational and therapist characteristics [e.g., Ref. (25, 36, 37)] as well as case-mix characteristics [e.g., Ref. (15)] have been associated with implementation outcomes.

Beginning in fiscal year 2010–2011 and within the context of a state budget shortfall, LACDMH-contracted agencies and directly operated programs were offered the opportunity for reimbursement for the delivery of a number of evidence-based and community-defined EBPs through the PEI transformation in children’s mental health services. LACDMH furnished initial implementation support (i.e., initial training and consultation) for six EBPs for children and adolescents, including CBITS, CPP, MAP, SS, TF-CBT, and Triple P, which were selected for supported implementation based on both the presenting problems (not diagnosis) targeted and the capacity of the EBP developers to train very large numbers of therapists within a short time frame (38). **Table 1** provides a brief summary of these EBPs. Training was ongoing throughout the study time frame; therapists could be trained and begin claiming under PEI at any time within the 19 fiscal quarters. Funding support for PEI training and delivery was also ongoing, as dictated by the California Mental Health Services Act, which was passed by voters in 2004. This is a permanent state-funding stream that can only be terminated or altered by a majority of state voters through a new ballot initiative. However, county plans for fund administration may be subject to change.

**Procedures**

The current study extracted administrative PEI claims data for the six initial EBPs supported by LACDMH for implementation in children’s mental health services, spanning 19 fiscal quarters, or approximately 57 months, between fiscal years 2009–2010 (Quarter 4) and fiscal years 2014–2015 (Quarter 2). These data capture the initial rollout through early sustainment period of this EBP implementation effort. Claims for this study were restricted to “psychotherapy” units of service that were delivered to clients under 21 years old (defined as youth by LACDMH), that occurred between May 11, 2010, and December 31, 2014, and that were delivered by therapists who billed at least three psychotherapy claims during this time frame. These data represent 6,873 unique therapists who were employed within 88 unique agencies and billed a total of 2,322,389 psychotherapy claims. Claims were aggregated to the therapist level for the delivery of each practice and for the delivery of any of the six EBPs.

As part of the larger 4KEEPS Project (27), this study was approved by the Institutional Review Board at the University of California, Los Angeles, CA, USA.

**Participants**

A total of 6,873 therapists were represented in the extracted claims across the study period. Therapist demographic, professional, and case-mix characteristics derived from the claims data are provided in **Table 2**.

**Measures**

All therapist characteristics were derived from the claims data. For each categorical variable, the largest category was selected as the reference group. The following therapist demographics were included as categorical predictors in each model: primary language (English, Spanish, other), discipline/type at the time of therapist’s first PEI claim [marriage and family therapist (MFT), rehabilitation professional, counselor, social worker, trainee, psychiatrist, other], and race/ethnicity (Hispanic/Latino, non-Hispanic White, other non-Hispanic minority).

Therapist service characteristics included the following continuous predictors: the average number of claims that the
TABLE 1 | Indicated age range, target problems, and consultation and training requirements for the six EBPs as noted in the PEI Implementation Handbook, Revised July 2016.

<table>
<thead>
<tr>
<th>Indicated age range (years)</th>
<th>Target problems</th>
<th>Ongoing consultation</th>
<th>Minimum training required before claiming is allowed</th>
<th>Train-the-trainer allowed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Behavioral Intervention for Trauma in Schools</td>
<td>10–15</td>
<td>PTSD, traumatic stress</td>
<td>Weekly consultation calls for at least 10 weeks are required</td>
<td>2-day on-site</td>
</tr>
<tr>
<td>Child–Parent Psychotherapy</td>
<td>0–6</td>
<td>Trauma, poor attachment</td>
<td>Bi-weekly group consultation calls for 18 months; 6- and 12-month booster trainings are required</td>
<td>Initial 2½ days</td>
</tr>
<tr>
<td>Managing and Adapting Practice</td>
<td>0–23</td>
<td>Anxiety, conduct, depression, traumatic stress</td>
<td>Twice-monthly consultation calls for 6 months are required (unless trained by an agency-based supervisor)</td>
<td>8 h</td>
</tr>
<tr>
<td>Seeking Safety</td>
<td>13+</td>
<td>PTSD, substance use</td>
<td>Consultation calls are not required</td>
<td>Initial 6 h</td>
</tr>
<tr>
<td>Trauma-Focused Cognitive Behavioral Therapy</td>
<td>3–18</td>
<td>PTSD, traumatic stress</td>
<td>12 consultation calls and a booster training are required</td>
<td>Webinar and initial 2-day in-person</td>
</tr>
<tr>
<td>Positive Parenting Program</td>
<td>0–18</td>
<td>Disruptive behavior, family dysfunction</td>
<td>Consultation calls are not required</td>
<td>Initial training (1–3 days)</td>
</tr>
</tbody>
</table>

therapist billed to PEI per active day, the average number of unique clients served per month, the total number of agencies at which the therapist claimed for one of the six PEI EBPs during the time frame, and the total number of EBPs out of six for which a therapist claimed. An active day was defined as a day in which a therapist made at least one claim. The practice for which each therapist made the most claims was included as a predictor in the model examining continued delivery of any of the six EBPs.

Therapist case-mix characteristics included the following continuous predictors, which were determined based on the percentage composition of a therapist's total caseload or total claims during the time frame: client admission diagnosis (percentage of a therapist's caseload whose admission diagnosis was an adjustment disorder or a disorder other than mood/anxiety, disruptive behaviors, ADHD, or trauma), client ethnicity (percentage of a therapist's caseload that is Hispanic), and service setting (percentage of a therapist's claims that take place in an office rather than in field settings (home, school, other community locations)). In addition, the average client age and client gender (percentage of a therapist's caseload that is male) were included in the model as continuous variables. We examined the percentage of a therapist's caseload whose primary diagnosis was an adjustment disorder or other disorder than mood/anxiety, disruptive behaviors, ADHD, or trauma) was included as an indicator of agency size.

The mean and median lengths of delivery (i.e., survival times) were calculated for the delivery of each practice and for the delivery of any of the six EBPs. Kaplan–Meier (KM) survival functions were generated as well, and differences across the six EBPs were determined using the log-rank, Wilcoxon, and Tarone–Ware tests of survival function equality. Since the results of the three tests did not differ, only results from the log-rank test are reported below.

Factors Associated with Risk of Discontinuation of Any EBP

A multivariate Cox regression (semi-parametric survival analyses) model was performed to determine the unique contribution of each predictor variable to the sustainment of therapists’ overall EBP practice delivery. The Cox regression model was selected because, as a semi-parametric model, no assumption had to be made about the distribution of the survival time (40). Survival time represented the time elapsed, in units of months, from the time of the therapist's first claim to the time of the therapist's final claim for any of the six EBPs. The binary outcome variable was (1) sustained delivery (i.e., censored) vs. (2) discontinued delivery (i.e., failure event). Sustained delivery was right-censored and defined as a continued claiming through the end of available claiming data, which was the fourth fiscal quarter (Q4), or the final 3 months, of 2014, between October 1 and December 31, 2014. Discontinued delivery was defined as not claiming for any of the six EBPs during this final quarter of our data. Therapists in fiscal years 2011–2012 and 2012–2013 (38). Whether or not an agency has multiple sites can also be construed as a binary indicator of agency size.
who “paused” claiming for EBPs over a 3-month period prior to 2014 (e.g., in 2012, 2013, or 2014) but who resumed claims were not considered to experience a discontinuation event.

A single model examining the predictors of the delivery of any of the six EBPs was conducted. The model controlled for whether the therapist began billing for PEI services during the initial rollout period in 2010 (i.e., early entry) or 2011 or later (i.e., later entry). In addition, consistent with the “shared frailty” approach used by Aarons et al. (29), we included agency ID for the last agency at which a therapist claimed as a term in the model, in order to account for the unobserved agency-level random effect, or shared frailty, of therapists nested within an agency (40). Therapists working at the same agency are presumably subject to the same external environment (e.g., agency climate), which suggests that therapists of a single agency would have a “shared” or a “common” value for their frailty, which represents the therapist’s inherent but unmeasured likelihood of experiencing the event of interest (i.e., discontinued delivery) (40). For 88.9% of the therapists, the final agency was the only agency at which the therapist claimed.

Following a test of the proportional hazards assumption of Cox regressions, a few variables violated the assumption of proportionality, which were consequently entered into the model as having time-varying coefficients: the average number of claims made per active day, the number of agencies at which therapists billed to PEI for one of the six EBPs, and the number of EBPs for which therapists billed to PEI. For all categorical variables, the category represented by the largest number of therapists was selected as the reference category.

All analyses were performed using Stata/SE 13.0 (41).

## RESULTS

On average, therapists made 337.9 (SD = 467.09) claims to the six EBPs of interest and delivered these interventions to 22.25 (SD = 28) clients across the 57 months under study. In this sample, 6,111 (88.9%) therapists made psychotherapy claims for at least one of the six EBPs at only one agency, 652 (9.5%) billed at two agencies, 89 (1.3%) at three agencies, 19 (0.3%) at four agencies, and 2 (0.03%) billed at five agencies. Therapists claimed for an average of 2.18 (SD = 1.11) EBPs (range = 1–6) during the time frame of our data. In addition, 2,387 (34.7%) therapists claimed for one practice; 29.7% claimed for two EBPs, 21.1% claimed for three EBPs, 12.1% claimed for four EBPs, and 2.4% claimed for five or six EBPs. Two thousand and thirty-seven therapists (29.6%) made their first PEI claim for these six EBPs in 2010, whereas 1,651 (24.0%), 1,411 (20.5%), 1,068 (15.5%), and 706 (10.3%) therapists began claiming in 2011, 2012, 2013, and 2014, respectively. One hundred and thirty-nine therapists (2.0%) made their final PEI claim for these six EBPs during the time frame of our data. Based on making their first PEI claim on or before December 31, 2010, 2,037 (29.6%) therapists were in the initial cohort of therapists involved at the outset of this system-driven implementation effort.

Table 2 shows the distribution of therapists who made the first PEI claim for each of the six EBPs in 2010, 2011, 2012, and 2013. Therapists of a single agency would have a “shared” or a “common” value for their frailty, which represents the therapist’s inherent but unmeasured likelihood of experiencing the event of interest (i.e., discontinued delivery) (40). For 88.9% of the therapists, the final agency was the only agency at which the therapist claimed.

On average, therapists made 337.9 (SD = 467.09) claims to the six EBPs of interest and delivered these interventions to 22.25 (SD = 28) clients across the 57 months under study. In this sample, 6,111 (88.9%) therapists made psychotherapy claims for at least one of the six EBPs at only one agency, 652 (9.5%) billed at two agencies, 89 (1.3%) at three agencies, 19 (0.3%) at four agencies, and 2 (0.03%) billed at five agencies. Therapists claimed for an average of 2.18 (SD = 1.11) EBPs (range = 1–6) during the time frame of our data. In addition, 2,387 (34.7%) therapists claimed for one practice; 29.7% claimed for two EBPs, 21.1% claimed for three EBPs, 12.1% claimed for four EBPs, and 2.4% claimed for five or six EBPs. Two thousand and thirty-seven therapists (29.6%) made their first PEI claim for these six EBPs in 2010, whereas 1,651 (24.0%), 1,411 (20.5%), 1,068 (15.5%), and 706 (10.3%) therapists began claiming in 2011, 2012, 2013, and 2014, respectively. One hundred and thirty-nine therapists (2.0%) made their final PEI claim for these six EBPs during the time frame of our data. Based on making their first PEI claim on or before December 31, 2010, 2,037 (29.6%) therapists were in the initial cohort of therapists involved at the outset of this system-driven implementation effort.

### Table 2: Therapist-level demographic, service, case-mix, and agency characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Categorical variables</th>
<th>Continuous variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Made first claim in 2010 (early entry control)</td>
<td>2,037</td>
<td>29.6</td>
</tr>
<tr>
<td>Discipline/type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marriage and family therapist</td>
<td>1,954</td>
<td>28.4</td>
</tr>
<tr>
<td>Rehabilitation professional</td>
<td>1,407</td>
<td>20.5</td>
</tr>
<tr>
<td>Counselor</td>
<td>1,355</td>
<td>19.7</td>
</tr>
<tr>
<td>Social worker</td>
<td>795</td>
<td>11.6</td>
</tr>
<tr>
<td>Trainee</td>
<td>530</td>
<td>7.7</td>
</tr>
<tr>
<td>Other</td>
<td>648</td>
<td>4.4</td>
</tr>
<tr>
<td>Psychiatrist</td>
<td>184</td>
<td>2.7</td>
</tr>
<tr>
<td><strong>Primary language</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>3,968</td>
<td>56.3</td>
</tr>
<tr>
<td>Spanish</td>
<td>2,392</td>
<td>34.8</td>
</tr>
<tr>
<td>Other</td>
<td>613</td>
<td>8.9</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>2,392</td>
<td>34.8</td>
</tr>
<tr>
<td>Other non-Hispanic minority</td>
<td>2,422</td>
<td>35.3</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>2,059</td>
<td>30.0</td>
</tr>
<tr>
<td><strong>Service characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average number of EBPs claims made per active day</td>
<td>1.81</td>
<td>0.83</td>
</tr>
<tr>
<td>Average number of clients served per month with EBP</td>
<td>1.38</td>
<td>2.54</td>
</tr>
<tr>
<td>Number of agencies at which therapists billed</td>
<td>1.13</td>
<td>0.40</td>
</tr>
<tr>
<td>Number of EBPs for which therapist billed</td>
<td>2.18</td>
<td>1.11</td>
</tr>
<tr>
<td><strong>Case-mix characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client race/ethnicity (% of a therapist’s caseload)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>64.13</td>
<td>29.31</td>
</tr>
<tr>
<td>Other non-Hispanic minority</td>
<td>22.65</td>
<td>25.01</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>9.40</td>
<td>15.51</td>
</tr>
<tr>
<td>Client primary presenting problem/admission diagnosis (% of a therapist’s caseload)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalizing disorders: mood or anxiety disorders</td>
<td>42.56</td>
<td>26.75</td>
</tr>
<tr>
<td>Externalizing disorders: disruptive behavior disorders or ADHD</td>
<td>29.94</td>
<td>24.26</td>
</tr>
<tr>
<td>Adjustment or other disorders</td>
<td>17.56</td>
<td>22.40</td>
</tr>
<tr>
<td>Trauma disorders</td>
<td>9.09</td>
<td>16.06</td>
</tr>
<tr>
<td>Client average age</td>
<td>11.79</td>
<td>3.40</td>
</tr>
<tr>
<td>Client gender (% males)</td>
<td>53.66</td>
<td>27.13</td>
</tr>
<tr>
<td>Service setting (% of a therapist’s total claims)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office (outpatient)</td>
<td>57.28</td>
<td>37.32</td>
</tr>
<tr>
<td>Home</td>
<td>19.78</td>
<td>27.02</td>
</tr>
<tr>
<td>School</td>
<td>13.12</td>
<td>21.36</td>
</tr>
<tr>
<td>Other</td>
<td>9.07</td>
<td>20.97</td>
</tr>
<tr>
<td><strong>Agency characteristic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final claim agency having multiple sites</td>
<td>5,158</td>
<td>75.0</td>
</tr>
</tbody>
</table>
Characterizing Continued EBP Delivery

Among all therapists, 2,443 (35.5%) continued delivery of any of the six EBPs at the end of the study time frame. Table 3 displays the mean and median survival times, as well as the frequency of discontinued delivery for each practice. Figure 1 shows graphical illustrations of the KM survival functions for therapist delivery of each practice and of any of the six EBPs of interest.

A visual inspection of Figure 1 indicates that CBITS had a higher risk of discontinuation than the delivery of the other EBPs. The log-rank test of survival function equality revealed significant differences across the six EBPs, $\chi^2 = 207.1$, df = 5, $p < 0.001$. To identify the specific EBPs that were different from the rest, follow-up log-rank tests were performed to compare the survival curve of each practice to the combined survival curve of the five other EBPs (29). Results revealed that the survival curve for CPP delivery did not significantly differ from the survival curve of the other EBPs ($\chi^2 = 0.02$, df = 1, $p = 0.878$).

Factors Associated with Risk of Discontinuation of Any EBP

Table 4 and Figure 2 display the results of the multivariate model including predictors of discontinued delivery of any of the six EBPs. For further ease of interpretation, please refer to Figure 2 for illustration of the relative risk of significant predictors. After controlling for whether therapists made their first claim in 2010 or later, results revealed a number of variables to be significantly associated with a risk of discontinued delivery.

TABLE 3 | The mean and median survival times (months) for EBP delivery in the descending order of median survival time.

<table>
<thead>
<tr>
<th></th>
<th>Total therapists</th>
<th>Events* (n, %)</th>
<th>Censoredb (n, %)</th>
<th>Mean survival time (months)</th>
<th>Median survival time (months)</th>
<th>Min (months)</th>
<th>Max (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any of the six practices</td>
<td>6,873</td>
<td>4,430 (64.5%)</td>
<td>2,443 (35.5%)</td>
<td>21.73</td>
<td>18.70</td>
<td>0.033</td>
<td>56.10</td>
</tr>
<tr>
<td>Managing and Adapting Practices</td>
<td>4,328</td>
<td>2,830 (65.4%)</td>
<td>1,498 (34.6%)</td>
<td>18.16</td>
<td>15.30</td>
<td>0.033</td>
<td>53.80</td>
</tr>
<tr>
<td>Trauma Focused-Cognitive Behavior Therapy</td>
<td>4,392</td>
<td>3,239 (73.7%)</td>
<td>1,153 (26.3%)</td>
<td>18.78</td>
<td>14.60</td>
<td>0.033</td>
<td>55.77</td>
</tr>
<tr>
<td>Child–Parent Psychotherapy</td>
<td>950</td>
<td>666 (70.1%)</td>
<td>284 (29.9%)</td>
<td>16.53</td>
<td>12.60</td>
<td>0.033</td>
<td>56.10</td>
</tr>
<tr>
<td>Positive Parenting Program</td>
<td>1,807</td>
<td>1,479 (81.8%)</td>
<td>328 (18.2%)</td>
<td>16.18</td>
<td>11.43</td>
<td>0.033</td>
<td>54.77</td>
</tr>
<tr>
<td>Seeking Safety</td>
<td>3,353</td>
<td>2,447 (73.0%)</td>
<td>906 (27.0%)</td>
<td>16.58</td>
<td>11.20</td>
<td>0.033</td>
<td>55.63</td>
</tr>
<tr>
<td>Cognitive Behavioral Interventions for Trauma in Schools</td>
<td>145</td>
<td>142 (97.9%)</td>
<td>3 (2.1%)</td>
<td>8.90</td>
<td>3.97</td>
<td>0.467</td>
<td>50.87</td>
</tr>
</tbody>
</table>

*The number of events represents the number of therapists who discontinued delivery (i.e., made no claims during the final fiscal quarter of 2014).

bThe percentage of total therapists who continued to deliver during the final fiscal quarter of 2014.

CBITS ($\chi^2 = 97.84$, df = 1, $p < 0.001$), SS ($\chi^2 = 14.81$, df = 1, $p < 0.001$), and Triple P ($\chi^2 = 58.89$, df = 1, $p < 0.001$) had a significantly higher risk of delivery discontinuation than the other EBPs, whereas MAP ($\chi^2 = 60.86$, df = 1, $p < 0.001$) and TF-CBT ($\chi^2 = 4.05$, df = 1, $p < 0.05$) had a significantly lower risk than the other EBPs. These results align with the patterns visible in Figure 1.

![FIGURE 1](cumulative_survival.png) | Cumulative Kaplan–Meier survival functions for therapist delivery of each EBP and of any of the six EBPs.
Later entry control (reference = early entry) & 1.681*** & 0.060  
**Therapist demographics**  
Therapist type/discipline (reference = marriage and family therapist) &  
Counselor & 1.241*** & 0.057  
Social worker & 1.126* & 0.059  
Rehabilitation professional & 1.124* & 0.052  
Psychiatrist & 1.708*** & 0.149  
Trainee & 1.941*** & 0.129  
Other (e.g., Case Manager, Psychologist, etc.) & 1.234*** & 0.072  
Therapist primary language (reference = English) &  
Spanish & 0.904* & 0.041  
Other & 1.032 & 0.055  
Therapist ethnicity (reference = Non-Hispanic White) &  
Hispanic & 0.950 & 0.047  
Other non-Hispanic minority & 0.902** & 0.034  
**Therapist service characteristics**  
Average number of claims made per active day & 0.983*** & 0.001  
Average number of clients served per month & 1.179*** & 0.007  
Number of agencies at which therapists billed to PEI & 0.996* & 0.002  
Number of evidence-based practices for which therapists billed to PEI & 0.987*** & <0.001  
Practice for which therapist made the most claims (reference = Trauma-Focused Cognitive Behavior Therapy) & 2.104** & 0.51  
Cognitive Behavioral Interventions for Trauma in Schools & 0.640*** & 0.059  
Child–Parent Psychotherapy & 0.693*** & 0.027  
Managing and Adapting Practice & 0.832* & 0.045  
Seeking Safety & 0.813** & 0.057  
Positive Parenting Program &  
**Case-mix characteristics**  
Client ethnicity (% of caseload that is Hispanic) & 0.999 & <0.001  
Client admission diagnosis (% of caseload that is adjustment or other disorder) & 0.999 & <0.001  
Client average age & 1.022** & 0.007  
Client gender (% of caseload that is male) & 1.001 & <0.001  
Service setting (% of claims that occurred in the office) & 1.0001*** & <0.001  
**Agency characteristic**  
Final agency having multiple sites (reference = single site) & 1.141*** & 0.042  
  
HR, hazard ratio.  
*p < 0.05, **p < 0.01, ***p < 0.001.

practice delivery. Note that, for categorical variables, hazard ratios indicate how high the risk of discontinuing delivery is for a therapist in one group compared to a therapist in another group, if all other variables were held constant. For continuous variables, hazard ratio indicates a change in the risk of discontinuing delivery if the variable/predictor of interest is increased by one unit (42). For example, the hazard ratio of 0.984 for the average number of daily claims indicates that, for each additional claim made per active day and holding other variables constant, the risk of discontinuing delivery is 0.984 times lower (or 1.6% lower) than the risk of discontinuing delivery for a therapist who makes one fewer claim per active day. In the same example, for a therapist who makes 10 more claims per active day—and holding all other variables constant—the relative risk is (0.984)^10 = 0.851, or a 14.9% lower risk of discontinuing delivery.

**Therapist Demographic Characteristics**  
Counselors, social workers, rehabilitation professionals, psychiatrists, trainees, and therapists of other disciplines were at 24.1, 12.6, 12.4, 70.8, 94.1, and 23.4% higher risk, respectively, of discontinuing practice delivery than were MFTs. Therapists with Spanish as their primary language exhibited a 9.6% lower risk of discontinuing delivery than therapists whose primary language was English. In addition, therapists who identified as other non-Hispanic Minority demonstrated a 9.8% lower risk of discontinuing practice delivery, compared to non-Hispanic White therapists.

**Therapist Service Characteristics**  
Each additional claim made per active day was associated with a 1.7% decreased risk of discontinuing delivery of any EBP. Each additional unique client served per month was associated with a 17.9% increased risk of discontinuing practice delivery. Each additional agency at which a therapist claimed was associated with a 0.4% decreased risk of discontinuing practice delivery. Each additional EBP claimed for in total across the study time frame was associated with a 1.3% decreased risk of discontinuing practice delivery. Therapists who most frequently billed for CPP, MAP, SS, and TP exhibited 36, 30.7, 16.8, and 18.7% lower risk, respectively, of discontinuing delivery of any of the EBPs, compared to therapists who primarily billed for TF-CBT. Therapists who primarily billed for CBITS exhibited a 110.4% higher risk of discontinuing any EBP delivery, compared to those who primarily billed for TF-CBT.

**Case-Mix Characteristics**  
With respect to case-mix composition, neither the percentage of a therapist’s caseload that is Hispanic nor the percentage of a therapist’s caseload that presented with an adjustment or other disorder was significantly associated with an increased or a decreased risk. The average age of a therapist’s clients was significantly associated with the therapist’s risk of discontinuing practice delivery; each additional year of clients’ average age was associated with a 2.2% increased risk. The proportion of male clients on a therapist’s PEI caseload was not significantly associated with a risk of discontinuing practice delivery. With respect to service setting, the percentage of a therapist’s total claims that were made in the office was associated with a 0.01% increased risk of discontinuing practice delivery.

**Agency Characteristic**  
A therapist whose final claim was made at a multisite agency exhibited a 14.1% higher risk of discontinuing delivery than a therapist whose final claim was made at a single-site agency.

**DISCUSSION**  
This study highlights a novel application of survival analysis to understand EBP sustainment using administrative claims data to track system-level sustainment of six EBPs over 19 fiscal quarters. Administrative claims were made by therapists delivering EBPs in the context of a system-driven, fiscally mandated implementation
of EBPs (i.e., the PEI transformation); PEI funding was available throughout the study time frame. Results revealed that the average survival time for any of the six EBPs within the 57-month study time frame was 21 months, with the average survival time for individual EBPs differing significantly with a range from 9 (CBITS) to 19 months (TF-CBT). Overall, therapist demographic, case-mix, and service characteristics, as well as agency characteristics (centralization), were significantly associated with a risk of therapists’ discontinuation of any EBP. Consequently, these conditions may have implications for return on investment in EBP training.

The first aim of this study was to characterize sustained delivery of any of the six EBPs and examine differences by EBP. As shown in Table 3, the mean and median survival times for the
delivery of each EBP or of any of the six EBPs were under 2 years, suggesting that our 5 years of claims data allow for interpretable conclusions. As would be expected, the median survival time is lower than the mean, in part because of the presence of positive outlier therapists who have continued to bill for long periods. However, the mean survival time can be tricky to interpret when there are unequal observation times for each therapist; for example, a therapist could have begun delivering an EBP with only a few months of observation time remaining, with a substantial portion of censored data. The mean is therefore dependent on the time frame of this specific study, and the mean time of actual usage in the field is likely to be even longer than what is reported. The median survival time is less susceptible to the influence of study time frame and is more reflective of the median time of actual usage.

Relative to a combined survival curve of the delivery of five other EBPs, therapists who primarily delivered CBITS, SS, and Triple P had a higher risk of discontinuation, whereas therapists who primarily claimed for MAP and TF-CBT had a lower risk of discontinuation within the system. It is unsurprising that CBITS exhibited such a high risk for discontinued delivery; it was never adopted widely by therapists, in part due to its limitation as a school-only EBP (38). Indeed, studies have found program leaders to have positive perceptions of MAP due to the wide range of cases or clients that MAP can be used with (39, 43). In addition, MAP, TF-CBT, and CPP require ongoing consultation, which may have implications for their sustained delivery (44). By contrast, SS and TP do not require ongoing consultation.

The current findings are somewhat consistent with findings on volume-based penetration using the same dataset, in which Triple P, CPP, and CBITS had a lower volume of claims, relative to MAP, TF-CBT, and SS (15). However, the present study found that SS had a significantly higher risk of delivery discontinuation than the other EBPs, and that CPP risk did not differ significantly from the other EBPs. These differences reflect how the analysis of different types of sustainment outcomes (claims volume/penetration vs. therapist discontinuation) may generate both convergent and divergent findings.

The second aim of this study was to identify factors associated with the likelihood of sustained practice delivery for any of the six EBPs. Our model controlled for whether a therapist started claiming for PEI in the first year of PEI rollout or later. Starting with workforce characteristics as predictors, social workers, trainees, psychiatrists, counselors, therapists of other disciplines (e.g., case managers, psychologists), and rehabilitation professionals at the time of their first claim were all more likely to discontinue delivery than MFTs. Particularly striking are the hazard ratios for trainees and psychiatrists, who exhibit nearly twice as much risk of discontinuing delivery (94.1 and 70.8%) as MFTs. These findings suggest that allocating EBP training resources—at least for these six EBPs—toward temporary employees may represent shorter-term investments.

We found that therapists whose primary language was Spanish were at a significantly lower risk of EBP discontinuation. This group represents 34.8% of the therapist workforce represented in the data. The finding suggests that therapists who are prepared to serve the large proportion of non-English, Spanish speakers in the County system are retained in the EBP delivery workforce. The results indicate that efforts to recruit bilingual, bicultural mental health providers may provide excellent returns on EBP-training investments in individuals who are best able to reach typically underserved populations.

Not surprisingly, making more claims per day (i.e., greater volume of therapist PEI claims) was associated with a decreased risk of discontinuing delivery; however, serving more unique clients per month was associated with an increased risk of discontinuation. These somewhat contrary findings may relate to therapist burnout. Indeed, past research has shown that a high caseload is associated with an increased burnout (e.g., Ref. (29, 45)). An increased number of unique clients controlling for the number of units of service delivered daily may translate to increased requirements for documentation and outcome monitoring with each additional unique client served. By contrast, an overall higher volume of EBP delivery may facilitate greater mastery that contributes to a longer continued use by therapists.

Therapists who billed for more EBPs or at multiple agencies were at a lower risk of discontinuing delivery of any of the six EBPs. This finding is encouraging, as it suggests that even when workforce turnover occurs at the agency level, there may be recapture of EBP-training investments at the system level. Relative to therapists who made the most claims to TF-CBT, therapists who made the most claims to CPP, TP, MAP, or SS were all at a lower risk of discontinuing any EBP delivery. By contrast, therapists who primarily delivered CBITS exhibited a much higher risk (more than twice) than therapists who primarily delivered TF-CBT. The latter finding is unsurprising given the lower sustainment of CBITS relative to TF-CBT documented in our prior work (15).

With respect to client case-mix characteristics, therapists with a higher proportion of Hispanic clients were at a lower risk of discontinuing any EBP delivery when compared to therapists with a higher proportion of clients of other ethnicities. These findings suggest that when a given therapists’ caseload primarily resembles the most prevalent client profiles served in a given system (46) (i.e., younger, Hispanic/Latino children presenting with mood/anxiety disorders served in school settings), EBP delivery retention is more likely. Having older child clients on average was also associated with a higher risk for therapists to discontinue practice delivery. This may be explained by the fact that the coverage of the EBPs under study predominantly targets children rather than adolescents. We did not find an association between the risk of discontinuation and caseload representation of youth with admission diagnoses other than mood, anxiety, conduct, and trauma problems targeted by the six EBPs. However, the average representation of these problems was low on average in the sample.

With respect to organizational characteristics, the centralization of the agency and the primary service setting type were associated with a risk of delivery discontinuation. First, therapists at agencies where services were primarily school-based had longer tenures of sustaining EBP delivery compared to therapists at office-based sites. This finding may suggest that therapists primarily delivering care in settings with fewer access barriers may show more longevity in EBP implementation. Second, claiming at
a multisite agency was associated with a higher risk of discontinuing EBP delivery than being at a single-site agency. Centralized, single-site agencies tend to also be smaller than multisite agencies. Previous findings from the same system context have suggested that larger agencies installed more systematic strategies at multiple levels (i.e., organization, therapist, client) to support initial EBP implementation (38). However, the current findings may indicate that the greater resources put in place by larger agencies may not ensure EBP sustainment at the therapist level. However, this is in contrast to research demonstrating turnover to be lower where employees are more embedded in their job and work in larger organizations (47).

Some limitations of the present study must be noted. The limited time frame of current data precludes analysis of sustainment beyond 19 fiscal quarters; indeed, 35.5% of therapists continued to bill for one of the six EBPs within the final fiscal quarter of our analysis. However, survival analysis accounts for those therapists who are considered to be “censored” to produce a reliable model. In addition, these data do not shed light on whether therapists discontinued delivering PEI EBPs altogether or whether they might have initiated or continued to deliver PEI EBPs other than the six examined in this study. Relatedly, this study was only able to examine the sustainment of these six approved PEI practices, as they were the only ones initially selected by the LACDMH for implementation support (38). Furthermore, we were unable to track therapist migration to mental health agencies not billing to PEI; it is plausible that these therapists continued to deliver one of the six EBPs at another agency (e.g., a private practice, an agency outside of Los Angeles County) not represented in the LACDMH claims data. In addition, we were not able to control for other variables that have been associated with sustainment, such as community readiness, the extent to which program staff received support and assistance, and other contextual factors (e.g., Ref. (20, 48)). While we are unable to track specific instances of “turnover per se,” our finding that therapists who provided services at more than one agency had better survival odds indicated that turnover across agencies may not be inconsistent with therapist-level sustainment of EBP delivery in this context. A limitation inherent to using administrative claims data is that we infer “delivery” when the data itself only truly indicate “billing.” Claims data also do not indicate whether a practice was delivered with fidelity. In addition, importantly, the claims data included in the current study do not represent the entirety of a therapist’s practice, that is, these data only represent that therapist’s administrative claims for these six EBPs for children or transition-age youth. Therapists likely served other many other children through different funding sources and/or other EBPs covered under the PEI program, whereas some therapists may also have served clients of other age ranges.

Despite these limitations, this study has important implications for system-driven implementation efforts. This study illustrates the novel contribution of applying survival analysis methods to administrative claims data to examine returns on system-level investments in workforce training. The findings provide a benchmark for continued therapist EBP delivery within a system (vs. individual organizations). Furthermore, the findings suggest potential mutable factors to target in sustainment interventions. For example, the findings highlight that strategic assignment of therapists to EBP training should be based on maximizing fit between the EBP and the therapist’s existing case mix. Likewise, the findings also underscore the importance of relevance-mapping approaches to the system- and agency-level selection of EBPs for adoption with the goal of long-term sustainment (14).

**Author Contributions**

LB-F and AL contributed to the study design/framework, outlined the introduction, and drafted the discussion. CZ contributed to the study design, research methodology, performed the analyses, and drafted the introduction, methods, results, and discussion. NS contributed to the study aims and methods. DS, SR, and GA contributed to the research methodology and manuscript editing. DI-G and LB facilitated access to the data used in this study and contributed to the interpretation of findings. All authors have reviewed this manuscript.

**Funding**

This study was funded by a grant from the National Institute of Mental Health (R01MH100134).

**References**


Cleves MA, Gould WW, Marchenko YV. *Introduction to Survival Analysis Using Stata*. College Station, TX: StataCorp Corporation (2016).

StataCorp L. *Stata: Release 13-Statistical Software*. College Station, TX: StataCorp LP (2013).


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer KH and handling Editor declared their shared affiliation.

Copyright © 2018 Brookman-Frazee, Zhan, Stadnick, Sommerfeld, Aarons, Innes-Gomberg, Bando and Lau. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Agency Leaders’ Assessments of Feasibility and Desirability of Implementation of Evidence-Based Practices in Youth-Serving Organizations Using the Stages of Implementation Completion

Lawrence A. Palinkas¹, Mark Campbell² and Lisa Saldana²*

¹ Department of Children, Youth and Families, Suzanne Dworak-Peck School of Social Work, University of Southern California, Los Angeles, CA, United States; ² Oregon Social Learning Center, Eugene, OR, United States

Background: This study examined influences on the decisions of administrators of youth-serving organizations to initiate and proceed with implementation of an evidence-based practice (EBP).

Methods: Semi-structured interviews, developed using the Stages of Implementation Completion (SIC) as a framework, were conducted with 19 agency chief executive officers and program directors of 15 organizations serving children and adolescents.

Results: Agency leaders’ self-assessments of implementation feasibility and desirability prior to implementation (Pre-implementation) were influenced by intervention affordability, feasibility, requirements, validity, reliability, relevance, cost savings, positive outcomes, and adequacy of information; availability of funding, support from sources external to the agency, and adequacy of technical assistance; and staff availability and attitudes toward innovation in general and EBPs in particular, organizational capacity, fit between the EBP and agency mission and capacity, prior experience with implementation, experience with seeking evidence, and developing consensus. Assessments during the Implementation phase included intervention flexibility and requirements; availability of funding, adequacy of training and technical assistance, and getting sufficient and appropriate referrals; and staffing and implementing with fidelity. Assessments during the Sustainment phase included intervention costs and benefits; availability of funding, support from sources outside of the agency, and need for the EBP; and the fit between the EBP and the agency mission.

Discussion: The results point to opportunities for using agency leader models to develop strategies to facilitate implementation of evidence-based and innovative practices for children and adolescents. The SIC provides a standardized framework for guiding agency leader self-assessments of implementation.

Keywords: innovation, adoption, feasibility, desirability, evidence-based treatments and practices, qualitative methods, youth mental health, SIC
INTRODUCTION

In the past two decades, there has been an increased effort to implement evidence-based practices (EBPs) into real world community public social service settings (1, 2). To do so successfully, agency leaders are necessarily involved in extensive planning and self-assessments of the organization’s capacity to engage in training and quality assurance in the EBP, which often involves a complex set of interactions among developers and system leaders, front line staff, and consumers (3). In fact, it is generally understood that it takes an agency a minimum of 2 years to complete the full implementation process (4) for psychosocial interventions and that the success of a program is strongly influenced by the presence or absence of certain barriers and facilitators to implementation (5–8) as well as the strategies selected to facilitate the implementation (9–11). Yet, there remains much to be learned regarding which aspects of these methods and interactions are most valuable for successful installation of new practices (8) and which are considered by agency leaders when conducting self-assessments throughout the full implementation process.

There is consensus that implementation of psychosocial interventions within social service settings is a recursive process with well-defined stages or steps (5, 12). Fixsen and Blase (4) described several stages that are not necessarily linear and that impact each other in complex ways. In the case of EBPs, an intervention developer, technical assistance purveyor, or third-party intermediary typically assists programs in navigating their way through the implementation process to ensure that the program elements are delivered in the manner intended by the developers. Nevertheless, there remains a dearth of methods to accurately assess the key processes involved in the implementation stages and the fidelity of the implementation methods assessed (13, 14).

Given the non-linear, yet staged progression of the implementation process, the measurement of this process must be flexible enough to capture these potential variations. Having a well-defined measure for assessing implementation to inform knowledge about the typical progression through the stages of implementation helps to increase the likelihood that sites can use information garnered in the early stages to inform and support their success in later stages (5, 8). There might be particular value in giving agencies feedback regarding their progress during the early implementation stages to help them assess and potentially calibrate their efforts to proceed, or to reassess whether their current implementation plan remains viable (3, 15). The Stages of Implementation Completion [SIC; (3)] was developed to meet this need.

The Stages of Implementation Completion (SIC)
The SIC is an 8-stage assessment tool (12) developed as part of a large-scale randomized implementation trial that contrasted two methods of implementing Treatment Foster Care Oregon [TFCO (formerly Multidimensional Treatment Foster Care); (16)], an EBP for youth with serious behavioral problems in the congregate care and child welfare systems. The eight stages range from Engagement (Stage 1) with the developers/purveyors in the implementation process, to achievement of Competency in program delivery (Stage 8). The SIC was developed to measure a community or organization’s progress and milestones toward successful implementation of the TFCO model regardless of the implementation strategy utilized. Within each of the eight stages, sub activities are operationalized and completion of activities are monitored, along with the length of time taken to complete these activities.

As described in Methods, the current paper is part of a larger trial focused on adapting the SIC for additional EBPs within youth-serving systems, and assessing its utility in measuring implementation success across service sectors including schools, substance abuse treatment, and juvenile justice (3). Sub activities within each stage target specific tasks necessary to complete each stage within the particular EBP. For example, Readiness Planning (Stage 3) is required for the implementation of all EBPs, but one TFCO specific sub activity is to conduct a foster parent recruitment review. The SIC is date driven in order to analyze the pace of implementation, and to identify when agencies experience hurdles in their process that might delay success. The SIC also yields a proportion score which takes into account the number of activities within a stage that are completed. Thus, scores for both the speed and the proportion of activities are calculated to determine if such factors influence the successful adoption of an EBP. The SIC maps onto three phases of implementation including Pre-Implementation, Implementation, and Sustainment.

Agency Leader Considerations for Implementation

Although there are numerous frameworks and models for implementing healthcare innovations, each consistently point to several factors believed to be associated with successful adoption, implementation, and sustainment, including characteristics of the inner and outer setting, the intervention itself (i.e., characteristics of the organization and staff involved in the implementation and the external environmental factors that determine the demand for the intervention and the resources available for its implementation, respectively, and the population served by the practice (5–8). As Proctor et al. (17), p. 72) have observed, “the success of efforts to implement evidence-based treatment may rest on their congruence with the preferences and priorities of those who shape, deliver, and participate in care.” Similar pragmatic arguments have been made by Bhattacharyya et al. (18) who pose that theory is not necessarily better than “common sense” for guiding implementation. For example, highlighting the importance of fiscal viability and feasibility, a study by Palinkas et al. (19) found that directors of youth-serving mental health clinics made decisions to adopt or not adopt innovative EBPs based on an assessment of costs and benefits associated with adoption, capacity for adoption, and acceptability of new practices. Moreover, assessment of costs and benefits exhibited several principles of behavioral economics including loss aversion, temporal discounting, use of heuristics, sensitivity to monetary incentives, decision fatigue, framing,
and environmental influences. However, the extent to which these preferences and priorities change from one stage of implementation to the next (19) and how decision-makers self-assess the weight of these considerations remains unknown.

**Current Paper**

The current paper aims to disentangle these targets by integrating firsthand knowledge from agency leaders working within youth-serving systems when attempting to implement one of three EBPs for the treatment of serious emotional problems in youth. To better understand the key factors involved in deciding whether or not to adopt, support, and sustain one of these practices, we conducted a qualitative study focused on the assessment of feasibility and desirability of the EBP, as well as the efforts made for implementation by mental health and social service agencies representing different youth-serving systems. Our aim was to examine the factors that influenced these self-assessments made by agency leaders at each stage of implementation, as measured by the SIC.

**METHODS**

**The SIC Study**

The study described in this paper was part of a larger effort to determine if the SIC can be applied across EBPs from different service sectors and accurately predict successful implementation outcomes. Naturally occurring implementation attempts were observed toward program start-up and sustainment of three widely-used EBPs that operate in various contexts and child public service sectors: Multisystemic Therapy [MST; (20)], Multidimensional Family Therapy [MDFT; (21)], and TFCO (16). Selection criteria for inclusion in this study included being an EBP: (a) for child and family mental health delivered within key service sectors (schools, juvenile justice, substance use, child welfare); (b) large real-world uptake within the respective service sectors in order to achieve sample size requirements for study analyses; and (c) EBP developers who expressed interest in using the SIC and in advancing understanding of the universal implementation elements shown to increase the successful uptake of EBPs.

**Participants**

A random sample of 5 sites, from each of the three participating EBPs, that had advanced minimally through Stage 3 (Readiness Planning) or beyond were recruited for participation. Identified sites were asked to consent to participate in a one-time semi-structured qualitative interview conducted by phone. Participants were compensated ($45) for their time. Consenting procedures included a request for use of the sites’ SIC data being collected as part of the larger study. Of the 15 sites identified, 14 agreed to participate in this phase of the overall study.

A total of 19 individuals participated across the 14 agencies, with three agencies including two representatives, and one agency including three representatives. Agencies were asked to purposefully select individuals within their agency that were responsible for decision making regarding the program, and who were involved in the implementation planning. The majority of the participants (68.4%) were female, and ages ranged from 31 to 63 years. Participants included agency leaders (i.e., chief executive officers [CEO], executive directors, or deputy directors; 31.6%); program specific directors (42.1%); or clinical supervisors (26.3%). The study was approved by the Institutional Review Board at the Oregon Social Learning Center.

**Data Collection**

Semi-structured interviews were conducted over the telephone by two project investigators, either in tandem or separately, using a standardized interview guide organized using the SIC as a framework. Questions focused on behaviors and perceptions regarding implementation activities identified on the SIC (e.g., “What steps or processes did you go through at your agency before getting started to determine if the EBP would be a good fit for your agency?”). Information collected during the interviews included characteristics of person being interviewed, role in agency, involvement in implementing the EBP, the process of implementation at each of the SIC stages that were completed by the agency, potential for sustainment of the EBP, reflections on consistency of the EBP with participant and agency goals, perceived costs and benefits associated with implementing the EBP, primary changes in any of the existing policies and procedures that were necessary to implement the EBP, and anything in particular the participant or agency staff did or did not like about the EBP in general. Interviews lasted approximately 1 h and were audio recorded, with written notes taken and used to supplement recordings when the audio quality was poor.

**Data Analysis**

Interviews were transcribed verbatim. Transcripts were reviewed for accuracy by two members of the research team. Deductive and inductive thematic coding was employed in this analysis (22). Deductive codes were based on interview questions, and inductive codes were based on responses by agency leaders. Data were coded according to Saldaña (23) first- and second-cycle coding method. During the first coding cycle, three researchers independently reviewed and open-coded these raw materials. They then undertook “consensus coding” together, a process used to establish agreement and increase the rigor and validity of coding in qualitative research (24, 25). Lists of codes developed individually by each investigator were subsequently discussed, matched and integrated into a single codebook. Inter-rater reliability in the assignment of specific codes to specific transcript segments was assessed for a randomly selected transcript. For all coded text statements, the coders agreed on the codes 81% of the time, indicating good reliability in qualitative research (26). A web-based qualitative data management program (27) was used for coding and generating a series of categories arranged in a treelike structure connecting text segments as separate categories of codes or “nodes.” Through repeated comparisons of these categories with one another, these nodes and trees were used to create a taxonomy of themes that included both a priori and emergent categories and new, previously unrecognized categories. The SIC was used as a framework for organizing the
first order codes, while grounded theory analytic methods (28) were used to construct themes based on the inductive codes.

RESULTS

Analysis of the self-assessments made by system leaders in the course of adopting, supporting and sustaining the three EBPs revealed distinct influences on the decisions about the feasibility and desirability of the EBP and its implementation as measured by each of the three SIC phases—Pre-Implementation, Implementation, and Sustainment (Table 1). Each of these sets is described below with an emphasis on themes related to the intervention, outer (external factors such as policy and funding) and inner (internal factors such as staffing and attitudes) settings, and the interaction of these factors.

Pre-implementation Phase (SIC Stages 1–3)

Preliminary assessments of feasibility and desirability of the EBP and its implementation at the Pre-Implementation phase corresponded to characteristics of the intervention, the outer and inner settings common to many implementation frameworks (5–7).

Intervention

Characteristics of the intervention that system leaders considered included its affordability, flexibility, requirements, validity, reliability, relevance, evidence of positive outcomes, and potential cost savings. For instance, one EBP was selected by an agency because “it was also a bit more affordable” and because “[EBP] is a much more flexible type of therapy to implement. And by flexible, just, I mean that in every sense of the word” (program director). On the other hand, one of the agencies that eventually elected not to proceed with a different EBP, instead selected an intervention that they judged to be less intrusive and burdensome. This same EBP was indeed selected by a different agency though due to its rigor as stated by one agency CEO, “Well basically, we, you know, are juvenile justice and we wanted a program that was valid, and reliable, and tested on an urban juvenile justice population with the challenges that our kids and families face.”

Agency directors also evaluated EBPs during the Pre-Implementation phase for their potential to produce cost savings for the agency. As described by one director, “Well, feasibility goes hand in hand with money…that’s what motivates everyone when you’re in a non-profit is that, you know, if I have an investment in staff, which is money, and I want a cost saving as it relates to [saving] kids from escalating, then it’s about how can I get the best bang for my buck?” Cost savings, in turn, are related to outcomes: “So we needed something that was intensive but over a long enough duration, to get us the outcome that we wanted… and then we looked at cost savings, in so far as, you know, we looked at that we wanted to keep kids in their community vs. having kids go out of home to a residential placement” (CEO).

Outer Setting

Characteristics of the outer setting included availability of funding and support for implementation, external demand for the intervention, and interorganizational relationships. Availability of funding was the most frequently cited characteristic of the outer setting during the Pre-Implementation phase. As one agency director noted, “We were definitely sure that we thought it would be a good fit with our background in treatment foster care. The main question was just whether or not we could get some funding to start it up.”

A second characteristic of the outer setting was the availability of technical assistance from the treatment developers. As noted by one CEO, “I guess the thought process that helped, that really aided in deciding that we would implement [EBP] was all of the support … they were able to review sample policies, give me feedback on those things, before we ever needed to kind of hit the ground running…I would say that’s a large pro, and that is probably what helped us to [do] decide, okay, let’s do this. Because, otherwise, we may not have been so quick to move into implementation” (CEO).

A third characteristic of the outer setting was an external demand for the intervention. Agencies that worked with larger...
service systems such as child welfare and juvenile justice would often take into consideration the goals, preferences, and mandates imposed by these systems. For instance, “...the goal of Probation, uh, then and now is they really were looking for something that was, uh, working with families, working with these clients, and really preventing residential care. So, they knew it needed to be something, you know, intensive, something that really looked at, you know, multiple systems with the ultimate goal of keeping kids out of residential care or out of detention... so they were really looking at, ‘Can these models really provide that kind of outcome that we’re looking for?’” (supervisor).

Finally, agency assessment of feasibility and desirability were based on the nature of social interactions with external organizations and representatives, including potential collaborators and sources of financial support. For example, one program director noted, “So we met with them as well as with our mental health board and throughout our community just to be sure as we’re going into this process that we weren’t necessarily wasting our money and time that no one was going to use this service” Whereas other valuable interactions included intermediaries that provided assistance with program development. For example, “Sometimes, we needed kind of an objective person who wasn’t necessarily part of any of us to bring us together, more of, to appear to bring in the program. So definitely helping with the stakeholders, I guess, would be what I’m trying to say. So they assisted with stakeholder meetings. They also were very instrumental in helping us put together our risk assessment tools.”

Inner Setting

During Pre-Implementation, characteristics of the inner setting that agency leaders included in their self-assessment included staffing requirements, staff attitudes toward the EBP, organizational capacity, fit between the EBP and agency mission and organizational capacity, prior experience with implementing EBPs, experience seeking information about EBPs, and building consensus. Considerations of staffing patterns focused on whether current staff were available to implement the EBP or additional staff needed to be hired. In some instances, agencies felt they could implement with current staff, as noted by one supervisor, “We already had staff. You know what I mean? We, uh, didn’t have to do a bunch of recruiting people or anything like that.” In other instances, agencies perceived a need to hire additional staff and had to consider the feasibility of successfully doing so, “We looked at, uh, the level of staff that were required, and would we be able to tap into that level of staff here in our area or geographic area? Would we have to be looking outside of that? Would it even be possible to, uh—to get those—that level of staff on board?” (program director).

Directors also considered whether existing staff would be in support of changing their practice and adopting new interventions. In one agency that was not successful in its efforts to implement [EBP], a clinical supervisor explained: “So what happened was some of our team members were not quite ready. They were not as open to the ideas of new evidence-based practice, which is not working and a change and being different with clients than normal, you know?” In another agency, the clinical director stated: “I don’t think there were any major surprises, not to me any major surprises, but there were some philosophical ideas that [EBP] really focused on that were a stretch for some of the staff and administration.”

Similarly, the capacity of the agency to implement something new also was assessed in terms of availability of resources for training and staff retention. As noted by one executive director, “In those days we didn’t do a whole bunch of high-tech analysis. Basically it was coming from direct experience from successes and challenges that we were faced with by just running foster care, being a foster family agency. We knew we had a problem retaining our foster parents and training and supporting them...to be able to deal with the high-end youth challenges. We also knew we needed resources in that department in order to do the work better, and right.”

Another inner setting influence on assessment of feasibility during Pre-Implementation was the nature of social interactions within the agency. For the most part, the interactions were characterized by consensus built on shared values. According to one supervisor, “We did not have to do a lot of internal negotiation before we landed on [EBP]. I provided my clinical director with some of the other options and as we talked about it and reviewed them. We definitely came to believe this fit our agency much better.” An associate executive director of another agency stated: “We didn’t know enough to not be on the same page. As it came up, we were like ‘This is amazing. This is amazing. We have to have it to improve our quality of services and our outcomes.’ It was as easy as that.” Yet, in other instances, some level of negotiation was required before consensus was achieved. As described by one program director, “Well, we had to do some negotiations. For one, it’s a higher per diem than what they’re accustomed to. So some of that was about what all it entailed and helping them to truly see the difference between following that model in comparison to our current foster care, where it’s more of an as-needed basis.”

Interactions Across Settings

Importantly, responses by agency leaders highlighted that many of their identified influences transcended more than one setting. For instance, the degree of fit between EBPs and the agency’s mission reflects an interaction between characteristics of the intervention with the outer and inner settings. As explained by one agency program director, “We basically researched a few of them, and then, we kind of—I, myself and my executive director met and we tried to figure out which one would be most feasible for our population, our organization, what would fit in with our procedures, um, and was also enhancing our treatment and so on.” In another example, the program director of one agency reported working with the treatment developer to determine how to fit the requirements of the EBP with the capacity of the organization and delivery of the intervention in rural settings, “And then what happened was we would usually discuss that, you know; because there were a lot of requirements that we didn’t, we needed to make, to make our system fit those. And they didn’t always fit those. So we spent a lot of time just simply on the requirements of [EBP] and how could we make that work in these rural areas.”
A second illustration of the interaction between different influences across settings was the availability of information about the EBP and its implementation. Participants reflected a range of responses to the question of whether they had sufficient information about the EBP and its implementation to make an informed decision. In some instances, agency directors were satisfied with the information they received about the EBP and how to implement it. As stated by one program director, “[EBP] is incredibly thorough. And from the beginning, even before we had contracted with them, they sent us all of the materials and all of the information that we were gonna need. And then throughout the entire process, having been assigned [EBP purveyor] from the agency, we were on the phone with him very regularly, and as well as he was available to us really whenever questions arose… then they assigned an expert to us who remains with us now. Uhm, so we certainly never were at a lack for, ‘How does this work?’ We were given so much information, and the representatives from [EBP] were available to us at any time that we needed them.” Yet other supervisors noted, “having a little bit more information would have made a smoother transition.” Lack of information led to feelings of mistrust and lack of transparency and an underestimation of the time and effort required for successful implementation, “So, we were told it was gonna to be very simple, we could pull them out of treatment, we could meet with the families; we had no idea, um, that, or I had no idea that the trainer was going to be, like, a huge stickler for whatever it takes, go to the home. Like, we were not an agency that was going into the home. I was told we wouldn’t really have to go into the home. That very quickly changed… So, we were told the information but it wasn’t, I felt like it wasn’t completely transparent.”

Related to the availability of information was the agency’s experience seeking information. Many participants, for instance, reported conducting their own literature and internet searches for information about the proposed EBP. “I Googled and did research on [EBP], and then I did research, and then I did a Google search on evaluations of [EBP], downloaded those articles, read them, then I’d looked at, you know, how they have an evidence-based practice website. Looked at it then to— based on my population, what would be the best one to go to. [EBP] kept popping up, and so after a while, it kept coming up again and again and again, and then I said, ‘Okay cool.’” (CEO). Whereas, other participants reported getting advice from others, including treatment developers and usually found this information to be quite helpful in facilitating the decision of whether or not to proceed with implementing the EBP. “After we actually approached [EBP], and we had a series of conversations with him and they would do this like readiness to implement after conversations. When he started telling us all the do’s and don’ts, that is when we started saying ‘Whoa, whoa, may not be a good fit after all’” (CEO). Finally, directors sought information from agencies and other agencies that had implemented or were currently implementing the EBP. “We went to an organization…. for an agency that was already up and running doing [EBP]…so we went there and met with them. I personally met with their staff and their recruiter and it was a 2-day trip for that alone as well as then meetings we had with the mental health board locally and other community providers to present this as an option just to see if there was even buy-in.”

Implementation Phase (SIC Stages 4–7)
Similar to Pre-Implementation, self-assessments of feasibility and desirability of the EBP and its implementation at the Implementation phase corresponded to characteristics of the intervention, the outer setting, and the inner setting.

Intervention
Influences associated with the intervention itself included its flexibility and requirements for delivering the service. In some instances, perceived flexibility of the EBP requirements made the process relatively easy. As noted by one program director, “The staff that they had were ready, and that is the great part about [EBP]. What I love is that they can do it half-time or full-time, and then you can start out half-time, and work toward getting a full-time [EBP] therapist, which is what we’re currently doing.” In other instances, EBP requirements created challenges. The on-call requirement for [EBP] was especially challenging, either because of the lack of available staff or the lack of experience of staff being on call. One program director explained in contracting out to a private agency to deliver [EBP], “I went with a private agency that we had worked with for many years but they have been known to be very effective with substance use in this area, because they have the adult drug court contract; they have been doing it for 25 years. But in that, when you have a private setting like that, they have earned the right at that level to not be on-call all the time. They have also earned the right at that level to not be in-home, if they don’t want to. I do think that in-home would have, at least initially, would have been a better fit if we could have figured out a way in this area to do that. The on-call portion has definitely been a huge challenge.”

Outer Setting
Outer setting influences during the Implementation phase included the availability of funding to deliver the service, availability and adequacy of technical assistance, and availability of appropriate referrals. The limited availability of funding to provide the service to particular clients was cited by some agency directors as an implementation barrier at this phase. For instance, “Uh, the main referral problem we had would be we really serve Medicaid—kids with Medicaid and once in a while—a lot of the kids through Juvenile Justice have Medicaid, uhm, but once in a while there is a—a child that we just are not able to serve because of the Medicaid issue.”

A second important outer setting influence was the availability and adequacy of technical assistance. Most agencies appeared satisfied with the technical assistance received after the initial clinical training, noting high accessibility to an expert EBP consultant with frequent interactions. For instance, one agency director stated that “…compared to how we’re usually trained, absolutely; I think this was phenomenal training. And I think that it’s as good as it can be. You always think there are things you could do, follow-up training you know, 90, maybe 90 days after you’ve actually implemented … that’s what the EBP experts are there for.” Another director stated: “it was a really steep learning
curve. Uh, and I think it was manageable because we did have weekly consultations. I think that's the only thing that made it doable...” In fact, one program director of another agency stated a desire for even more face-to-face consultation after the initial training: “I mean, the only thing that we would've liked to be different, but we couldn't because of the location of where the trainer was, would be [that] we'd like more face-to-face contact.” Similarly, another director commented, “and one thing about the [EBP] training team...that came out and trained all of us and our other staff, they have always been incredibly helpful to us. They have made it absolutely wonderful and a really easy way to learn.

So even though it gets really intense, they're always terrific with us.”

In other instances, particularly in four agencies that ultimately did not succeed in implementing their EBP, the assessment of technical assistance and training was somewhat mixed at best. For instance, one agency director stated, “and I do think some of the staff through the training, um, didn't necessarily have the best experience with the trainer and I felt like there might have been some damage to their perception of...the effectiveness of the program because the trainer didn't necessarily do a good job getting buy in, um, you know getting the staff excited about a great program.” Furthermore, as explained by a clinical supervisor from the same agency, “you have 3 days where you’re inundated with information, you’re watching a lot of DVDs, you have the manual. But then it's like you leave and then it's like poof, go do it. So, for me, no. I don’t learn that way. Because I don’t feel like 3 days of training, onsite training and a manual, is ready... No. We didn’t feel ready at all.”

A third important outer setting influence on the self-assessment of feasibility and desirability during the Implementation phase, was receipt of appropriate referrals. Some agencies began to solicit referrals soon after training was completed. One agency program director reported soliciting referrals 2 weeks after training: “We have a direct referral source from Probation, and since Juvenile Probation is the one who asked us to do the treatment they already had some families for our court.” According to the program director of one agency, “It has not always been 100% smooth, you know. We’d be wanting to have a case, we’re ready for new cases 2 weeks from now, and it might take until 3 or 4 weeks before we get them. So, it's been a little bit, uh, challenging to maintain full client caseloads at all times.” For others, the period between training and referrals was much longer. The executive director of another agency reported that there were few referrals for the first 18 months after training and, “it was really slow. A lot of that was us having to go back to child welfare and having to educate them on what a good referral was. We had to spend a lot of time on that. And we still actually have to spend a lot of time on that.” Related was how to track and extend appropriate referrals. One agency elected to create a management system for tracking referrals, to “you know, [know] who goes where. Who are they seeing? Uhm, so I track that every week, and I update Probation on, you know, who has been referred, who are they working with, and that kind of stuff because Probation didn’t actually have like a—an existing system for that. So, we created one.”

Decisions also were based on relationships with referral sources. For example, one agency described having to educate the child welfare agency in their county, “because part of the problem is that in our county, the way the referrals come in, through child welfare services, we rely on the child welfare staff to really have the reunification plan and then know about our programs and understand it, and then make the referrals. Otherwise we don’t have any. And in the same token, to be able to educate our foster parents in terms of an option, actually create a stock of foster parents that are trained and ready in this way. Oh, that stakeholder meeting was very, very vital for us, and actually it still kind of is.” Referral barriers experienced by agencies during this phase of Implementation was another factor underlining these decisions. One agency noted “we had to get the type of licensure—one is family foster care, which anyone without any prior experience can apply for. Then there's the specialized license, which is required per our contract with the state. So I lose, probably I lose half right there.... Other barriers ... some folks are maybe interested in chronic services, already have other children in the home. They might already be foster parents and they have other children in the home, which we cannot place if there's already a placement like that in the home.... To be honest some of the folks that maybe we thought were going to be a great fit, like current foster families, actually haven’t been.”

Lack of community support was another barrier to ongoing Implementation. “We really did not want to end the service. It was more along the lines of we didn’t have the community support to keep it going. So without that, I guess the mental health board supporting us, or someone else within the community speaking on our behalf that this really is the best service that we could provide... without someone coming forward to agree to help us do that, it was really hard to sell it... I think that's where we really got kind of stuck, was, how do we convince these people that this is the direction that everybody should go in?”

Inner Setting

Influences associated with the inner setting were parallel to those identified in the self-assessment during Pre-Implementation and included the availability of existing staff or the need to hire new staff to administer the EBP with fidelity. As noted earlier, in some instances, agencies had existing staff who were trained to implement the EBP. As noted by one program director, "so, we've got two therapists that were already embedded into the agency doing other types of private and adult drug court work, and then they ended up, uh, they're just taking on... because the [EBP] is 6 months—4 to 6 clients every 6 months for us...” In other instances, additional staff were hired to implement the EBP. “The majority of our team was hired. So not only were we new to this agency, we were new to [EBP]. And we were all hired under the understanding that we would become certified [EBP] therapists.” In some instances, the need to hire additional staff was a consequence of high staff turnover. “I think our biggest internal problems have been turnover in our department.”

Similarly, the availability of experienced supervisors influenced Implementation progression. As explained by
one agency program director, “Uh, but what was also very beneficial to us, and we didn’t even really realize how beneficial it proved to be or would prove to be, [was] the supervisor that we hired for the program came to us from another [EBP] program…. She was new to the supervisory role in [EBP], but she had been doing [EBP] therapy for a couple of years. So, with the combination of her being available full-time for the new staff, the four new staff for whom [EBP] was brand-new, she was very experienced, and so that made it pretty seamless. Had we had somebody in a supervisory role who was also brand-new to the model, which I know many programs do, I think there would have been some different challenges… that would not have been as smooth.”

Interactions Across Setting
Finally, implementing the program with fidelity was a challenge encountered by the agencies during the Implementation phase that cross both the inner and outer settings. For some agencies, fidelity to the model did not fit with existing community standards for delivering services. As explained by one agency program director, “I think really what came from that is that the county was not willing to comply with the fidelity, for one. So the referrals that we were getting didn’t quite fit in with the model and we were losing foster parents left and right because we were not providing them with placements.” For other agencies, lack of capacity to document fidelity was a major challenge. As the associate executive director of one agency observed, “Uploading the videos has been a challenge because our IT infrastructure, for whatever reason, cannot handle the capacity to upload them. So we’ve had some problems with that but I think we’ve finally resolved that. It required our IT director to get involved.”

Sustainment Phase (SIC Stage 8)
The Sustainment Phase of the SIC does not include a full assessment of sustainability, but rather, if the agency is prepared to begin sustaining long-term (3). Assessment of the feasibility and desirability of sustaining the EBPs was influenced by a number of factors, including whether there were sufficient revenues to support the program, whether there was support from sources external to the agency, and whether there was a genuine need for the program. Further, whether the EBP was consistent with agency goals, and the costs and benefits of implementation impacted self-assessment of sustainment.

Intervention
Similar to Pre-Implementation, and Implementation phases, characteristics of the intervention influenced agency leader self-assessments of Sustainment. In many instances, agency leaders pointed to elements of EBPs that were not compensated. For instance, one program director stated, “That’s something we’re definitely working on… it’s encouraged, you know, even though they might not do it after hours. They do text and call with their clients during the daytime, all the time. So there is a lot of hours outside of just sitting, doing therapy, that are involved with [EBP]. That, you know, there is at least some, not full, but some compensation for that… So, yeah, but recovering the costs of the extra hours, it’s just a much more expensive type of therapy than your average.”

Inner Setting
Inner setting considerations for the self-assessment of sustainment included whether or not the EBP was consistent with agency goals. As noted by one program director, “So we were already big believers of I guess [EBP] because we—we always treated the family. What this gives us is another option, another type of service to treat the family.” Moreover, the decision to sustain an EBP involved an assessment of the costs and benefits of implementation to the agency. One of the major costs identified by providers at this stage of implementation is the reduced revenue due to more intensive services delivered to fewer clients. According to one program director, “We get a lot of our revenue from group therapies, we run a ton of groups for our adult model… And so, you know, clinicians who are doing [EBP] don’t run as many groups because they don’t have the time or the capacity to do so because [EBP] takes up so much of their time. And so…, they have less of a case load of course and you know, don’t run as many groups. So then that cuts into our revenue…”

On the other hand, agency leaders were able to identify benefits that could influence sustainment, “I think that it’s given—it’s opened our staff up to some new ideas and—and new ways of doing things, so I think that was definitely a benefit.” Similarly, another program director noted, “We’ve seen more benefits to it than—it’s more positive for us than negative. It’s a great system, it works really well with very specific families. We have really good outcomes. We’re grateful that we had an opportunity to be trained in [EBP] and to use the model. The biggest barrier is the fact that sometimes we don’t get reimbursed at the rate at which we would like to, and that it actually costs us to deliver the service.” Not surprisingly, this barrier carried more weight for agencies who were not achieving strong clinical outcomes, “The agency is committed to funding it, so the funding for the program I feel pretty secure about. But I don’t know if in another year or 2 years if I don’t see improvements in our outcomes and our outcomes matching more of what [EBP] research outcomes say, I can’t say that I would be able to justify the ongoing cost… if I don’t start seeing the outcomes in, you know, the next year to 18 months. I’m realistic in that I know we just started this, just completed the training and it is going to take the staff some time to really grow into the model. But I would like to see some improvements in our outcomes in the next year to 18 months or we are going to have to take a real serious look at is the cost, the cost of the program to justify the outcomes we are getting.”

Outer Setting
Related to sufficient revenues to support the program is a consideration of the stability of that financial support from external sources. “The revenue source that, uh, is paying for the bulk of this here is something called… basically 60% of the revenue comes from federal and [state] dollars, and the other 40% then the state covers so, that we are… our cost recovery is a rate-based system. We worked with the consortium to set that
rate. They in turn—we bill—they bill [the state]...you’re getting about 55–60% of your costs, but because of the funding that we have here in [different state], we are able to recoup 100%. Part of it comes from the feds, the other part, the state kicks in so that we can be reimbursed fully.”

Another influence associated with the outer setting was whether there was a genuine need for the program. In all instances, participants expressed a need for evidence-based approaches to treatment in general and to the three participating EBPs in particular. “I saw that that was a big need. Our agency, with the youths that we work with... we had done individual and group therapy for years, with different evidence-based practices, but there was just a component that was missing, and it was definitely the family part. Uh, we’d had parenting classes but we needed much more than that.”

Interactions Across Settings
Not surprisingly, agency leaders’ self-assessment of the potential to sustain the programs involved an interaction of both inner and outer setting characteristics. As noted by one program director, “I just reiterate that these are the higher level cases that typically if you didn’t have something like this, and I would say, even go as far as to say if you didn’t have this, then you would probably see kids fall through the cracks. They would go on to become, you know, at an adult level in the prison system, because these are kids that nobody else really, uh, knows how to help, or what to do with. And so, you know, it prevents that happening...them falling through the cracks. And it helps kids that, really, the outcomes usually are very bleak, have some better outcomes...So you are preventing all costs to the state in terms of the, you know, that detention cost for long term detention. And you are also preventing the youth going on to be adult, uh, involved with adult crimes.”

Finally, although fiscal barriers were noted across all three implementation phases, agency leaders were able to recognize that once the Sustainment phase was achieved, the impact of this barrier might be reduced through the interaction of inner and outer settings. “Wow, you know the initial training cost is the barrier. That’s the biggest, the biggest cost. The ongoing costs, our costs for next year our continued certification is going to be like $8,000. Really in the scheme of things, that really isn’t insurmountable at all...It really doesn’t affect our revenue. The staffing costs are gonna be the same whether we do [EBP] or not, which is our biggest cost. Um, so anything additional... it really doesn’t affect their productivity because their productivity with [EBP] and what we were asking from them before was not inconsistent.”

DISCUSSION
Qualitative interviews from this study supported the overarching premise that the SIC could accurately guide agency leaders in a self-assessment of the pre-implementation, implementation, and sustainment phases (Table 1). Responses from agency administrators indicated that activities identified on the SIC can accurately distinguish sites that proceeded with EBP implementation, and those that determined that progression was not appropriate for their agency. Extent of prior experience with implementing EBPs appeared to be a factor in implementation success: Four of the 15 agencies had no prior experience with using EBPs and three of these four agencies discontinued implementation of the selected EBP. Moreover, the SIC was able to identify implementation activities that, when asked about, highlight challenges and facilitators that contribute to the success, or not, of implementation efforts. Importantly, although the selected EBPs are similar in requirements and structure for program delivery (e.g., team approach, community-based, family treatment), it is illuminating that the SIC was able to provide a generalized framework that is applicable across interventions.

In this study, assessment of the feasibility and desirability of implementation of EBPs was found to involve different sets of influences: Those that occur during Pre-Implementation, Implementation, or at the beginning of Sustainment. All three phases revealed characteristics of the intervention, the inner and outer settings, and the interaction between settings. Pre-Implementation was influenced by intervention characteristics of affordability, feasibility, requirements, validity, reliability, relevance, cost savings, positive outcomes, and adequacy of information. Pre-Implementation outer setting characteristics included availability of funding, support from sources external to the agency, and adequacy of technical assistance. Inner setting characteristics of staff availability and attitudes toward innovation in general and EBPs in particular, organizational capacity, the fit between the EBP and agency mission and capacity, prior experience with implementation, experience with seeking evidence, and developing consensus. Self-assessments that occurred during the Implementation phase included intervention characteristics of flexibility and requirements, outer setting characteristics of availability of funding, adequacy of training and technical assistance, and getting sufficient and appropriate referrals; and inner setting characteristics of staffing and implementing with fidelity. During the Sustainment phase, assessments included intervention costs and benefits, outer setting characteristics of availability of funding, support from sources outside the agency, and need for the EBP; and the inner setting characteristic of the fit between the EBP and the agency mission.

The results offer four specific insights as to how agencies assess the feasibility and desirability of EBP implementation. First, consistent with the observation made by (5), different variables or influences might play crucial roles at different points of the implementation process. Availability of funding to support the EBP was a characteristic of the outer setting that influenced assessment of feasibility and desirability at all three implementation phases. EBP flexibility and requirements, adequacy of technical assistance from the treatment developer, and availability of qualified staff were important influences during the Pre-Implementation and Implementation phases but not the Sustainment phase. Assessment of costs and benefits of the EBP, support from sources external to the agency, need or demand for the EBP, and fit between the
EBP and agency mission were influences on assessment of feasibility and desirability at the Pre-Implementation and Sustainment phases but not the Implementation phases. Also, the number of influences appeared to have grown smaller with each subsequent implementation phase (20 to 7 to 5), thereby suggesting that agency leaders weigh more considerations for continued implementation efforts during Pre-Implementation, than once the program has launched and is underway.

Second, the results highlight continuity of particular influences across all three implementation phases. Availability of funding was influential at all three phases. Assessment of whether there was sufficient information necessary to implement the EBP during the implementation phase was based on the information accessed and provided during Pre-Implementation. Assessment of the costs and benefits of the EBP conducted during the Sustainment phase was based on the experience during the Implementation phase.

Third, the results suggest that influences do not operate independently but in combination with one another. For instance, the inner setting degree of fit between different EBPs and the mission of the agency was an influence of assessment of feasibility and desirability that reflected an interaction between characteristics of the intervention (i.e., relevance) with the outer (i.e., demand for EBP) and inner settings. The availability of information about the EBP and its implementation was embedded in characteristics of the EBP itself as well as the outer (i.e., adequacy of technical assistance from treatment developer) and inner setting (experience with seeking evidence and information). The existence of such combinations suggests the need to examine potential mediation and moderation effects when identifying predictors of agency assessment of feasibility and desirability. For instance, evidence of EBP cost-effectiveness might impact level of support from sources outside the agency.

Finally, the assessments of EBP implementation feasibility and desirability are based on different forms of engagement, including engagement with other EBPs at the Pre-Implementation phase and engagement with information or evidence, and with other stakeholders at the Pre-Implementation and Implementation phases. An earlier study by Palinkas et al. (19) reported that personal experience was an important source of “evidence” used by systems leaders in deciding whether or not to implement EBPs. Clinical experience is one of the types of evidence used in the practice of evidence-based medicine (29). Use of research evidence was found in an earlier study to be significantly associated with the final stage achieved, as measured by the SIC, of TFCO by county-level youth-serving systems in California and Ohio (30). Engagement with other stakeholders both within the agency and external to the agency suggests that implementation is a “trans-relational” phenomenon involving interactions with other agencies (31–33), researchers (34, 35), and intermediaries (36).

**Limitations**

This study has several limitations. As a qualitative investigation, the generalizability of these findings is limited to a sample of senior administrators of agencies serving children and adolescents. The specific needs and perspectives of this stakeholder group on assessment of EBP implementation feasibility and desirability will likely differ from those of other stakeholders. Surveys of a random sample of different stakeholder groups would increase the generalizability of these results. Further, the assessment was based on the implementation of three specific EBPs. It is unclear whether the findings could be generalized to other EBPs. Finally, we did not conduct follow-up interviews with study participants, thus limiting our ability to establish a causal linkage between assessment of feasibility and acceptability and potential influences.

**CONCLUSIONS**

Despite the limited scope of this qualitative evaluation, our results support the conclusion that the relevance of implementation domains identified by most implementation models and frameworks vary by phase of implementation. Some of the influences on assessment of feasibility and desirability transcend more than one phase, while other influences appear to operate in combination with one another. Future research will consider if there is congruence between the quantitative data collected via the SIC, and the qualitative perspectives of agency leaders in their implementation process. Such evaluations will allow us to better assess and guide agencies toward informed decision-making and successful implementation.

**ETHICS STATEMENT**

This study was carried out in accordance with the recommendations of the Institutional Review Board at the Oregon Social Learning Center. The protocol was approved by the OSLC IRB. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

**AUTHOR CONTRIBUTIONS**

All authors contributed this article. LS is the Principal Investigator of the study. She conducted some of the qualitative interviews and engaged in manuscript preparation. MC is a research associate on the grant, conducted all of the qualitative interviews, transcription validation, and coding. LP is a co-investigator on the study and conducted the qualitative analyses, summarization, and manuscript preparation.

**FUNDING**

This study and manuscript were supported by the National Institute of Mental Health R01 MH097748 and the National Institute on Drug Abuse R01 DA044745.
ACKNOWLEDGMENTS

The authors would like to acknowledge Katie Bennett for her assistance with coding and reliability coding and Caroline Dennis for her assistance with manuscript preparation. We are grateful for our partnerships with MST, MDFT, and TFCO developers and purveyors. We thank the agency leaders who participated in our interviews and shared their experiences.

REFERENCES

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Palinkas, Campbell and Saldana. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Advantages of publishing in Frontiers

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPEN ACCESS</td>
<td>Articles are free to read for greatest visibility and readership</td>
</tr>
<tr>
<td>FAST PUBLICATION</td>
<td>Around 90 days from submission to decision</td>
</tr>
<tr>
<td>HIGH QUALITY PEER-REVIEWS</td>
<td>Rigorous, collaborative, and constructive peer-review</td>
</tr>
<tr>
<td>TRANSPARENT PEER-REVIEWS</td>
<td>Editors and reviewers acknowledged by name on published articles</td>
</tr>
<tr>
<td>REPRODUCIBILITY OF RESEARCH</td>
<td>Support open data and methods to enhance research reproducibility</td>
</tr>
<tr>
<td>DIGITAL PUBLISHING</td>
<td>Articles designed for optimal readership across devices</td>
</tr>
<tr>
<td>IMPACT METRICS</td>
<td>Advanced article metrics track visibility across digital media</td>
</tr>
<tr>
<td>EXTENSIVE PROMOTION</td>
<td>Marketing and promotion of impactful research</td>
</tr>
<tr>
<td>LOOP RESEARCH NETWORK</td>
<td>Our network increases your article’s readership</td>
</tr>
</tbody>
</table>

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: [www.frontiersin.org](http://www.frontiersin.org)
Contact us: info@frontiersin.org | +41 21 510 17 00

FOLLOW US
@frontiersin