OPEN

Adapting Clinical Guidelines for the Digital Age

AMERICAN COLLEGE OF Medical Quality

An Evaluation Framework for a Novel Process to Codevelop Written and Computable Guidelines

Amrita Tailor, MPH[1], Susan J. Robinson, PhD, MS-HCI (Ret.)[1], Dyann M. Matson-Koffman, DrPH, MPH, CHES[1], Maria Michaels, MBA[1], Matthew M. Burton, MD[2], and Ira M. Lubin, PhD[1]; on behalf of the Workgroup on Evaluation of the Integrated Process

Integrated process evaluation workgroup members (listed alphabetically):
J. Rex Astles, PhD. (CDC), Matthew M. Burton, MD (Holistic Healthcare Solutions, LLC), Jennifer Clark (ASCP), Gema Dumitru, MD, MPH (CDC), Juliet Haarbauer-Krupa, PhD (CDC), Joel Harder, MBA (AiCPG), Ira M. Lubin, PhD (CDC), Priya Jakhmola (CDC), Dyann M. Matson-Koffman, MPH, CHES (CDC), Maria Michaels, MBA (CDC), Susan J. Robinson, PhD (CDC), Amrita Tailor, MPH (CDC)

Abstract

Clinical practice guidelines (CPGs) support individual and population health by translating new, evidence-based knowledge into recommendations for health practice. CPGs can be provided as computable, machine-readable guidelines that support the translation of recommendations into shareable, interoperable clinical decision support and other digital tools (eg, quality measures, case reports, care plans). Interdisciplinary collaboration among guideline developers and health information technology experts can facilitate the translation of written guidelines into computable ones. The benefits of interdisciplinary work include a focus on the needs of end-users who apply guidelines in practice through clinic decision support systems as part of the Centers for Disease Control and Prevention's (CDC's) Adapting Clinical Guidelines for the Digital Age (ACG) initiative, a group of interdisciplinary experts proposed a process to facilitate the codevelopment of written and computable CPGs, referred to as the "integrated process (IP)."[1] This paper presents a framework for evaluating the IP based on a combination of vetted evaluation models and expert opinions. This framework combines 3 types of evaluations: process, product, and outcomes. These evaluations assess the value of interdisciplinary expert collaboration in carrying out the IP, the quality, usefulness, timeliness, and acceptance of the guideline, and the guideline's health impact, respectively. A case study is presented that illustrates application of the framework.

Key words:

evaluation, clinical practice guidelines, computable guidelines, guideline implementation, evaluation framework

Introduction

Translating clinical and public health knowledge for timely and actionable health care decisions is essential to achieving the best health outcomes.[2] Clinical practice guidelines (CPGs) are a major avenue for informing health care professionals of recommended practices.[3] However, there are challenges in the ways CPGs are developed and implemented such as the amount in time it takes to develop and publish CPGs and their timely adoption in health care practice.[4] In 2018, a group of multidisciplinary experts convened to take on the challenges and offer solutions to bridge these gaps as part of the Centers for Disease Control and Prevention's (CDC's) Adapting Clinical Guidelines for the Digital Age (ACG) initiative.[5] At this meeting, a multiyear initiative began to create a health IT standard for developing computable guidelines and the *integrated process* (IP), an Agile way to codevelop written (eg, narrative or textual documents in paper-based or web-based publications) and computable guidelines.[1,6] Computable guidelines are created by translating CPGs into machine-readable or computer-interpretable formats, thereby allowing for easier implementation into clinical workflow.[7,8]

[1]Centers for Disease Control and Prevention (CDC)
[2]Holistic Healthcare Solutions, LLC

SDC Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.ajmqonline.com).

American College of Medical Quality 2023 Vol. 38(5S) S35–S45

DOI: 10.1097/JMQ.0000000000000140

Guideline development typically does not include a structured evaluation as part of the development process or implementation.[9–11] Evaluation is important in determining how well this new IP works to improve timely production and implementation of quality written and computable guidelines.

The IP requires frequent collaboration among guideline developers, informaticians (eg, knowledge engineers and software developers), end-users (eg, clinicians, local health IT staff, electronic medical record system vendors, and others involved in the implementation), health communication experts, and evaluators.[1] The IP is designed to ensure the guideline's format and content accelerate clinical decision support systems (CDSS) and other health information technology (IT) tools (eg, quality measures, case reports, and care plans).[1,12] A structured evaluation of the IP offers the means to increase multidisciplinary interactions and generate quality intermediate products to speed the workflow and produce better health outcomes.

In the IP approach, the written guideline version follows traditional forms of guideline methods in summarizing the evidence review and recommendations.[1] Unique to the IP method, the computable version of the guideline is developed at the same time as the written version. The computable version is more highly structured, so it can be read and interpreted by a health IT system (machine-readable), such as an electronic health record. It presents the guideline as a set of definitions, codes, and logic expressions.[6,13] While the intended outcome for many written guideline development efforts is dissemination with less focus on implementation support, the IP includes development and integration of computable guidelines within clinical processes.

A central feature of this new IP is the incorporation of Agile project management practices for guideline development. Agile methodologies help produce value for customers through collaborative, iterative steps of product development that give team members and end-users plenty of say in product design.[6] Although this requires additional planning and coordination during guideline development, the payoff is a product that meets users' needs and reduces the time and effort to put the recommendation into practice.

Boxwala and colleagues describe levels of knowledge that provide a representation of the guideline recommendations in various formats (ie, narrative or "L1," semistructured or "L2," structured or "L3," executable or "L4").[13] Although traditionally implemented sequentially, the IP describes the codevelopment of the written and computable guidelines such that more than one knowledge level is developed simultaneously and iteratively, allowing for mutual feedback between the authors of the written and computable versions with input from implementers.

The IP is represented as a cycle diagram in Figure 1, which outlines the 4 stages and 12 phases that reflect the codevelopment of written and computable guidelines and shows that evaluation occurs throughout the IP, not just at the end.[1] The codevelopment of these guidelines and evaluation occurs during:

- concept planning,
- drafting of the written recommendations (L1) and semistructured logic constructs (L2),
- completion of written recommendations (L1) and computer-interpreted structured logic (L3),
- development and implementation of executable code (L4).

The IP also includes communication and evaluation activities, the latter to determine to what extent process steps were followed, whether products (eg, the computable guideline) were useful, and whether intended outcomes were achieved. This includes evaluating the effectiveness of multidisciplinary collaboration (process), the inclusion of user feedback in the various forms of the guideline (product), and the reach and impact of the final products (outcomes). This process, product, and outcomes approach provides the foundation of the evaluation framework for the IP. Undertaking the IP within an evaluation framework provides the means to detect and correct deviations from the workflow to keep the IP on track and ensure useful products and the desired clinical outcomes are achieved.

Methods

Multiple methods were used to develop the IP's evaluation framework. An evaluation workgroup was established in 2018 as part of CDC's ACG initiative to support this development process.[7] This workgroup included guideline developers, informaticians, implementers, communicators, and evaluators. Input on the evaluation framework was also received from subject matter experts in the other workgroups of the ACG initiative. Evaluation models and standards suitable to evaluation framework development were identified through discussions with evaluation experts in public health, clinical guidelines implementation, and informatics. Brief literature searches were also conducted on contemporary relevant evaluation standards and models.

Questions in the evaluation framework were based on the activities in the 12-phase IP and a review of established standards for developing and implementing evidence-based guidelines.[1,3,12] These standards informed evaluation questions related to whether a

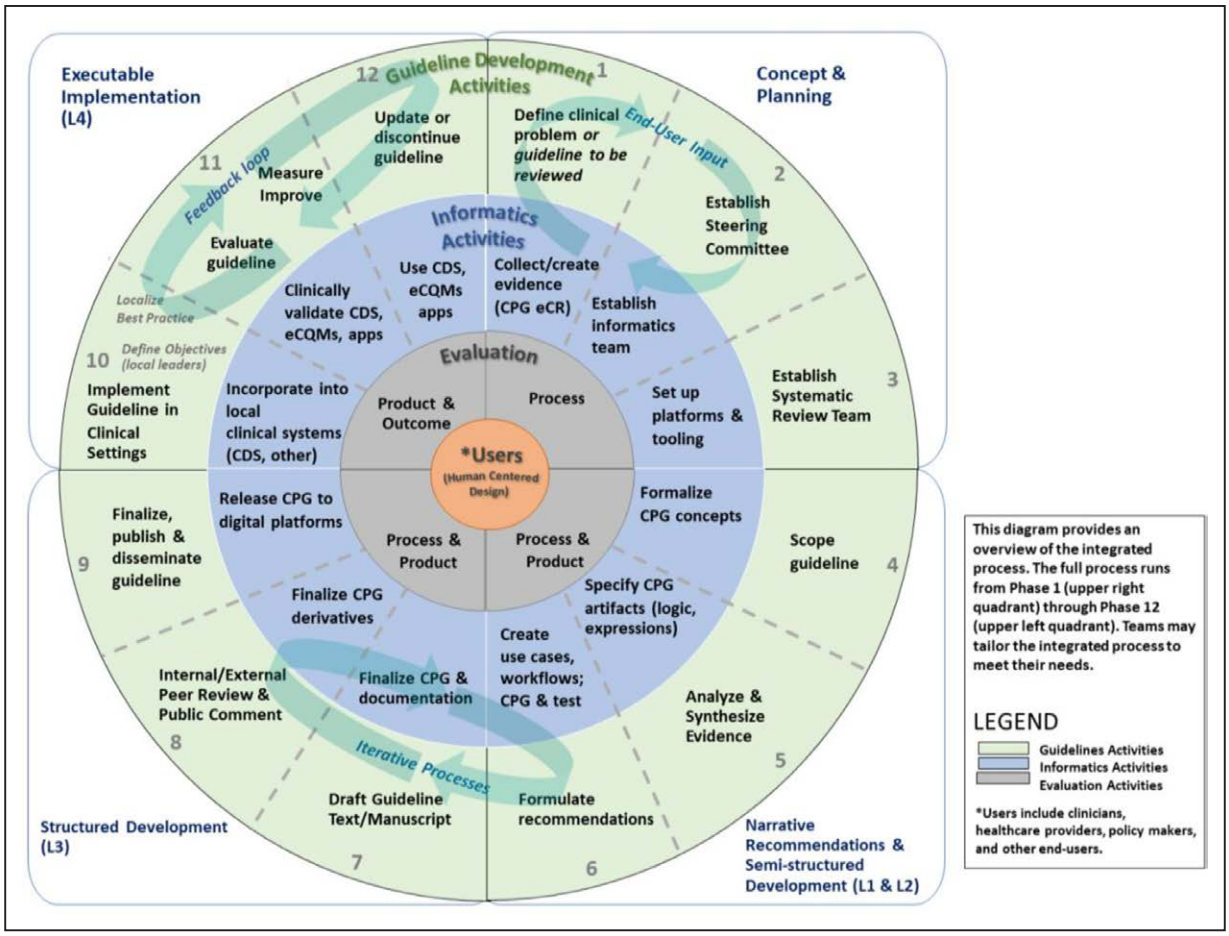

Figure 1. The 12-phase integrated process for guideline development and implementation. Abbreviations: CDS, clinical decision support; CPG, clinical practice guideline; eCQMs, electronic clinical quality measures; eCR, electronic case reporting.

scientific and rigorous evidence review process was followed, whether there was transparent reporting of how the body of evidence was linked to the strength of the recommendations (eg, the AGREE II tool),[14] and whether there was an appropriate translation of recommendations into computable guidelines for clinical implementation. An exploration of several theories and models such as the Diffusion of Innovation,[15] Actor-Network Theory,[16] Normalization Process Theory,[17] and general literature about implementation of health information systems[18–20] informed our understanding of how to implement recommended interventions aimed at becoming routine practice. To map out the detailed constructs of the IP evaluation framework, we relied on models that were aligned with our goals to assess the interaction between recommended interventions, external factors that influence their uptake, and characteristics of the human users—all within the context of evaluation: Human-Organization-Technology fit (HOT-fit),[21,22] Consolidated Framework for Implementation Research (CFIR),[23] and CDC's Framework for Program Evaluation.[13,24]

The HOT-fit framework focuses on three dimensions: *human* (system use and user satisfaction), *organization* (structure and environment), *technology* (system quality, information quality, and service quality).[21] This framework assesses both the individual influence of each of the dimensions and their interaction, thereby influencing implementation. For example, how characteristics of technology can influence human use and how the organizational environment also plays a role in this system.

On the other hand, the CFIR framework is broad and evaluates: the *innovation* (intervention) *characteristics*, *outer setting* (factors outside the organization), *inner setting* (factors within the organization), *characteristics of the individuals* (ie, end-users), and *process* (of implementation).[23] This framework is holistic in that it considers features of the intervention, external influence on implementation, self-efficacy measures of the end-users of the intervention,

environment within the facility, and engagement of relevant partners to optimize implementation. The HOT-fit and CFIR were useful in mapping out the evaluation questions for the *process* and *product* evaluation for the IP.

Finally, the CDC's Framework for Program Evaluation is a practical, flexible approach with six main steps*: engage collaborators, describe the program, focus the evaluation design* (ie, map the evaluation plan), *gather credible evidence, justify conclusions,* and *ensure use and share lessons*.[24] This design aims to help evaluators tailor their evaluation approach, identify the program's context, engage collaborators, and use the findings to improve the program. The concepts from this framework were useful in mapping out the evaluation questions related to the *process* and *outcome* evaluation.

Collectively, these frameworks, standards, and models, along with expert opinions, informed the evaluation questions for each phase of the IP and led to a comprehensive IP evaluation framework for assessing process, product, and outcomes.

Results

Suggestions from workgroup members, consultation with industry and government experts, the IP activities, and brief literature searches of standards and evaluation models informed the "12-Phase Integrated Process Evaluation Framework" (Excel Tool, available at https://stacks.cdc.gov/view/cdc/131007).[5] This evaluation framework is comprised of three main components: *process*, *product*, and *outcomes*. Each of these components focuses on specific aspects of the IP, as shown in the examples excerpted in Table 1.

The *process evaluation* asks: how is the guideline being created, disseminated, and implemented? The measurements of this evaluation component focus on the extent to which: (a) the team follows the steps outlined for codeveloping the written and computable guidelines; (b) there is early and ongoing engagement of experts in guideline development, informatics, implementation, communication, dissemination, and evaluation; and (c) the various disciplines of experts found the interaction with the other disciplines useful in developing the IP's intermediate and final products.

The *product evaluation* aims to determine if: (a) the written and computable guidelines are valid, reliable, and of good quality and (b) the guideline products are easy to use and useful to intended users. Ultimately, the products must be easy to implement and use in practice.

The *outcomes evaluation* examines whether the implemented guideline and related products achieve the expected short-term, intermediate, and long-term outcomes. These outcomes should be measured often after the guideline is released and implemented. The short-term outcomes include awareness and reach of the guideline among users. This can be measured soon after guideline dissemination (eg, 3–6 months after release). The intermediate outcomes focus on provider behavior change and policy changes that indicate organizational-level adoption of the guideline (eg, 6 months–1 year after release). Long-term outcomes assess improvement in health conditions, such as reductions in high blood pressure due to implementation of the guideline (eg, more than 1 year after release).

The amalgamation of the *process*, *product*, and *outcomes* components into a single evaluation framework is multifaceted. First, each component requires levels of assessment given the engagement of multiple disciplines during each phase of IP. Second, each component focuses on a different aspect of evaluation—process focuses on how the IP was executed, product focuses on what was produced, and outcome focuses on the impact of the products. The evaluation framework needs to include both the individual components and the interaction of these elements in one working environment. Third, the evaluation occurs over time, therefore, there may be evaluators for each of the components. To produce a robust and holistic evaluation, the process, product, and outcome components need to be compiled and summarized.

Evaluation Indicators of Process, Product, Outcomes Components

An evaluation indicator is a marker of progress and should be measurable.[24] Six evaluation indicators can be used to organize measures relevant to the IP's process, product, and outcome components in the evaluation framework, as described in Table 1.

The IP Evaluation Framework Tool

The *IP evaluation framework tool* is a practical resource for users to assess their guideline process, product, and outcomes components. The tool is flexible and can be tailored to each project or to stages of guideline development and implementation. Measures can also be selected to suit the needs of the user considering their priorities, time, and resources for the evaluation. The tool provides two trackers: an *overall tracker* and a combined *basic* and an *in-depth IP*

Table 1. Examples of Integrated Process Evaluation Components, Indicators, and Measures (*Overall Evaluation Tracker*)

Component	Overall Evaluation Questions	Indicators	Description and Measures
Process during all phases	**How was the guideline created and disseminated?** • To what extent were the phases and steps outlined in the integrated process followed to produce the guideline? • What are the perceived benefits and limitations in using the integrated process for guideline development and to support implementation at the local setting?	**Quality**	**The quality of the process used, eg, the integrated process** *Measures* 1. Steps that were followed and which were not 2. Descriptive accounts of benefits, barriers, and limitations, deviation from process
		Timeliness	**The time required to complete each of the phases, activities, major deliverables, and the overall integrated process** *Measures* 1. Data regarding timelines at end of each phase and at the mid/endpoint of the integrated process, (assess how far off from planned timeline & historical/other guidelines) 2. Factors that influenced timeline (ie, unanticipated factors) & lessons learned
		Resources	**The resources used during the integrated process (financial, human (expertise, time spent))** *Measures* 1. Reports of the financial and human resources (expertise, time spent) 2. Time permitting, resources used vs. resources allocated (during planning phases) to assess gaps, or resources that were needed but not added
Product During Phases 5 – 12	**Are the guidelines valid, reliable, and good quality, and how?** • Does use of the integrated process support the development of guidelines that are valid, reliable, good quality, and how? **Are the guidelines and related products easy to use (usable) and useful to the full range of intended users, and how?** • Did the guidelines and related products produced (using the integrated process) facilitate implementation of guidelines into clinical information systems?	**Quality**	**The quality of products or deliverables of the written and computable guideline** *Measures* 1. Quality of all major deliverables of written and computable guideline, looking at all individual quality assessment measures throughout the phases 2. Results of public comment and peer review 3. Results of user acceptance and usability testing 4. Descriptive accounts of benefits, barriers, and limitations of products during their implementation
Outcomes During Phases 10-12 Short-term (<6 months)	**Did the guideline and related products achieve short-term, intermediate, and long-term outcomes?** • Did the products derived from the integrated process facilitate achieving targeted health care objectives in a timely manner?	**Reach**	**The extent to which the guideline and related products were delivered to or requested by intended users and audiences** *Measures* 1. Number of web page visits and downloads 2. Number of downloads of clinic decision support code 3. User feedback to indicate awareness 4. Other measures of dissemination
Intermediate (6 months - 1 year)		**Behavioral and policy changes**	**Guideline use through provider behavior change and policy changes that indicate organizational-level adoption** *Measures* 1. Provider surveys to assess guideline use and changes in their behavior in accordance with the guideline 2. Usage metrics and data from guideline-based clinical decision support systems to assess provider utilization 3. Organizational leader surveys about any policy changes to indicate adoption of guideline across their organization
Long-term (>1 year after release)		**Health outcome improvements**	**Improvements in health outcomes or health impact due to implementation of the guideline** *Measures* 1. Guidelines-based clinic decision support utilization metrics that provide data on increases in screening rates and/or increases in the number of diagnoses 2. Surveys and/or focus groups to ask about earlier diagnoses or improvements in quality of life or health outcomes

evaluation tracker. The overall tracker requires synthesis of information *across* multiple phases (and steps), whereas the combined trackers ask about specific questions related to *each* phase and step of the IP. The basic and in-depth process evaluation trackers can be used to collect data to inform the

overall tracker related to the six indicators: *quality, timeliness, resources, reach, behavioral and policy changes, and health outcome improvements* (Table 1, available at https://stacks.cdc.gov/view/cdc/131007).[5] Users who have limited time or resources can direct their efforts to measuring select elements in the respective overall, basic, or in-depth evaluation tracker, depending on their focus.

Applying this tool is a subjective process to an extent. However, following the best practice of having a minimum of two individuals familiar with the guideline project to complete this tool and discuss their assessment will only strengthen the evaluation and credibility of findings compared to one person.

Overall Evaluation Tracker

This first part of the tool relates to the *overall evaluation* and allows users to document (1) the process of creating and disseminating the guideline (2) the guideline's validity, reliability, quality, and ease of use, and (3) whether the guideline and related products achieved specific outcomes. Responses to these elements are informed by measures related to the aforementioned six indicators (Table 1; available at https://stacks.cdc.gov/view/cdc/131007).[5]

Basic Evaluation Tracker

The *basic evaluation tracker* provides a mechanism for users to monitor progress on (1) whether each step in the IP was completed, (2) what type of individual expertise was involved and how it compares to what was recommended, (3) the level of usefulness of each respective step and involvement of relevant expertise to the overall development of the written and computable guidelines, and (4) challenges, solutions, and lessons learned during the implementation of each step (Table 2, available at https://stacks.cdc.gov/view/cdc/131007).[5] Users can track the progress of the process, product, and outcomes for each phase of the IP while identifying potential gaps in completing specific steps or phases. The tracker also serves as a way for users to record which steps were skipped, modified, or adapted. This documentation can be especially useful during future updates of the guideline.

In-depth Evaluation Tracker

The *in-depth evaluation tracker* (Table 3, available at https://stacks.cdc.gov/view/cdc/131007)[5] allows users to delve deeper into each IP phase to guide the assessment of the process, product, and outcome. These questions focus on areas related to the actions described or the quality of deliverables developed

Table 2. Example of the Basic Evaluation Tracker of Integrated Process: Phase 0, Steps 1-2

EXPERTISE INTEGRATION (BASIC EVALUATION) *Mark the expertise that was involved. The gray boxes note SMEs recommended in the integrated process*							INTEGRATED PROCESS		BASIC EVALUATION	
Org Lead or Management	GL Office or Lead	SME or Evidence Review Group	INFO	EVAL	COMM	IMPL	ID	Phases & Steps	Was this phase or step completed?	Rate the extent to which the phase or step was useful in developing the written and computable guidelines	Describe and list the challenges, solutions, lessons learned during this phase or step (process, product, outcome evaluation)
							0	**Pre-Guideline Development – Set up Operational Framework and an Integrated Process for Developing Written & Computable Guidelines**	Please Make a Selection	Please Make a Selection	
							0.1	**Identify functional and operational (including informatics and information technology) needs and requirements and limitations for developing written and computable guidelines in the organization.**	Please Make a Selection	Please Make a Selection	
							0.2	**Assess current resources and capacity for guideline development in the organization**	Please Make a Selection	Please Make a Selection	

COMM, Communication expert; EVAL, Evaluation expert; GL, Guideline; IMPL, Implementation expert; INFO, Informatics expert; SME, Subject Matter Expert.

Table 3. Example of the In-Depth Evaluation Tracker of the Integrated Process: Phase 0, Steps 1-2

INTEGRATED PROCESS		IN-DEPTH EVALUATION				
ID	**Phases & Steps**	**Type of Evaluation**	**Focused Questions**	**Responses**	**Additional Evaluator Notes/Comments**	**Additional Data Sources (e.g., document review)**
0	**Pre-Guideline Development – Set up Operational Framework and an Integrated Process for Developing Written & Computable Guidelines**					
0.1	**Identify functional and operational (including informatics and information technology) needs and requirements and limitations for developing written and computable guidelines in the organization.**	**Process**	What is the organization's prior experience with working on a written and computable guideline? How did it help or hinder in identifying relevant needs/requirements?			
0.2	**Assess current resources and capacity for guideline development in the organization**	**Process**	To what extent did the organization decide to include evaluation in carrying out the integrated process? [Why?]	Please Make a Selection		

during each step. Users also have the option of capturing comments or action items stemming from evaluation question responses to resolve issues in real time or to improve future projects. The in-depth evaluation tracker can help experts capture qualitative data about factors that influenced process, product, and outcome indicators (eg, length of time for development, resources needed, quality of the process or products, and reach of the guideline).

Case Study

One of the greatest challenges of guideline development and implementation is to gather sufficient resources to conduct evaluation during guideline creation and after guideline release. Even for large-scale efforts, it is not unusual that evaluation occurs after the guideline is already in use.[25] In this case study, we present how the evaluation framework and tool could be used for guidance development and implementation during the early months of the COVID-19 pandemic. This demonstration of the tool through an example case study is not intended to arrive at a formal conclusion. However, it shows how the evaluation framework and tool could have been used to streamline the collection of evaluation data during the COVID-19 response. The case study illustrates how the framework can help developers in the future, not only for traditional guideline development, but also for evidence-based guidance during public health emergencies.

In 2020, in response to the COVID-19 pandemic, a volunteer partnership of experts participating in the COVID-19 Healthcare Coalition (led by Mayo Clinic and MITRE Corporation),[26] formed the COVID-19 Digital Guideline Working Group (C19 DGWG), a subgroup of the broader coalition that included the American College of Emergency Physicians, the University of Minnesota, and other private industry partners.[27] Their goal was to rapidly develop a *COVID-19 Severity Classification and Disposition Recommendation Tool* to be used as guidance in triaging COVID-19 patients in emergency department (ED) settings. The triage guidance tool was scoped for clinical patient assessment along with severity classification and scoring based on clinical findings to inform treatment and disposition for COVID-19 patients in the ED. The final product was published and presented in the form of a visual flow diagram that could serve as a rapid and physical reference for health care professionals. This was coupled with a computable form of the tool that could be employed in various digital platforms, such as mobile applications, online calculators, or CDS.[28] The C19 DGWG used the Health

Level Seven International (HL7®) Fast Healthcare Interoperability Resources (FHIR®) Clinical Guidelines Implementation Guide (more commonly known as "CPG-on-FHIR®") as a standards-based framework for the development of the digital representation of the tool.[29]

The first evaluation step is to scope the efforts by selecting overarching evaluation questions followed by key analytic questions, corresponding indicators, and measures. For analyses after the guideline is developed or implemented, the scope is partly dependent on available data. For this case, the evaluator (SR, coauthor of this paper) selected the process question "How was the guideline created and disseminated?" and the product question "Are the guidelines valid, reliable, and good quality, and how?" Though these questions refer to a "guideline," the evaluator adapted them to assess the aforementioned "guidance tool." Conducting an outcomes evaluation was beyond the data and resources available. Due to interest in how well the IP worked in this emergency scenario, additional questions from the *Integrated Process Evaluation Framework Tool* were selected to examine (a) if the phases and steps in the IP were followed, (b) what challenges, solutions, lessons learned resulted from using the IP, and (c) how well the IP supported the production of the guidance tool's quality.

For the COVID-19 example case, data for evaluation came from various documents, intermediate work products generated during the project, published reports detailing the effort, and limited correspondence and discussions with participants involved in developing the tool.[30] To generate results, the evaluator used the evaluation tool to organize the data around relevant questions and measures for process and product evaluation. For example, the methodology described by C19 DGWG team members within CPG-on-FHIR® proved useful for answering process-related questions.[30] By mapping the methodology described in CPG-on-FHIR® to the IP steps detailed in the evaluation tool, one could see where the project workflow matched or did not match activities prescribed in the IP model. This mapping and review of other documents related to the development of the *COVID-19 Severity Classification and Disposition Recommendation Tool* generated a list of questions for the C19 DGWG co-leads to fill in gaps in what was known about the effort.

Using a combination of document review and brief interviews, the most important questions for each IP phase were relatively easy to answer during this informal demonstration. The team did not follow the IP in the organizing phases, *Pre-guideline Development (0)* and *Defining the Clinical Problem (1)* (see https://stacks.cdc.gov/view/cdc/131006),[5] due to the emergency nature of the COVID-19 problem. However, in latter phases (3–7, see https://stacks.cdc.gov/view/cdc/131006),[5] a review of activities showed an extremely high degree of integration among clinicians and informaticians. This was so much so that the term "Agile CPG Development"[6] was coined to describe the rapid codevelopment of written clinical concepts and recommendations together with computable elements. The people working together on the core clinical content and computable guidance were termed "Agile CPG Teams."[30]

The team members working on the guidance grappled with limited or emerging literature and practice-based experience being reported in real time due to pandemic conditions. This resulted in uncertainties in specifying clinical guidance and challenges for informaticians in developing concepts, data elements, and applying formal terminologies. Another challenge was the initial absence of formal terminologies for some of the concepts, as this was a novel disease. As a result, Step 6.2 (creating case presentations, use cases, user stories for recommendations where clinical decision support is appropriate, available at https://stacks.cdc.gov/view/cdc/131006 and https://stacks.cdc.gov/view/cdc/131007)[1,5] activities took on increased importance. The team relied heavily on case presentations, use cases, and user stories to synthesize the evolving evidence and hone written recommendations while applying the standards as best as possible in expressing those recommendations in computable form.

In developing the guidance tool in written and computable forms concurrently, the guidance team reviewed the evidence and produced successive drafts of the severity classification and recommendations in close collaboration with the informatics team. An early emphasis on translation to computable forms resulted in the guidance content being expressed in more formal and explicit artifacts (eg, Excel tables and diagrams), which quickly increased the specificity and usability of the recommendations. This helped maintain product quality for both the written and computable forms in a public health emergency. Thus, in Steps 6.3–6.6 (representing written recommendations into logical diagrams, mid-progress review, and designing, building, and testing CPG artifacts, see https://stacks.cdc.gov/view/cdc/131007),[5] product evaluation questions regarding accuracy of translation, quality, and performance of test cases were answered affirmatively, since the written and computable forms were consistent to a large extent. Over time, however, as the written guidance was updated, due to lack of available resources, the publicly available computable version did not continuously evolve as the scope of the original case study project did not include

updates. Further work on digital translations in the case study was accomplished by organizations and vendors focused on specific applications and users.

Use of the evaluation tool in this case study helped to examine what occurred throughout all phases and activities and revealed gaps in implementing the IP for this case. For example, despite an intent to recruit an evaluation lead, no one was available with evaluation expertise. The team similarly identified the need to have a dedicated communication expert to promote the effort and its products, but the all-volunteer team could not recruit one. While the C19 DGWG leads knew these disciplines were missing, it is possible that some of the evaluation and communication activities could have been included in the team workflows if this evaluation framework tool had been available at the start of the project. Due to the project's emergency nature, the implementation phases (phases 10–11, available at https://stacks.cdc.gov/view/cdc/131007)[5] were not formally tracked following the tool's release. However, the team continued to update the guidance tool as more evidence became available (phases 11–12, available at https://stacks.cdc.gov/view/cdc/131007).[5] A more thorough evaluation would have been necessary to determine what metrics could be generated to measure the guidance tool's effect on outcomes, though the CPG-on-FHIR® standard includes mechanisms to track specified outcome variables, such as clinical quality measures. Interviews with providers could have yielded insights into the usability, usefulness (to end-users), and utility (eg, functionality) of the guidance. This brief case study shows that the evaluation framework and tool can serve as a roadmap for gathering and organizing evaluation data during and after guideline or guidance projects.

Discussion

The evaluation framework is designed to inform the development of an evaluation plan tailored to the needs of guideline developers, informaticians, communicators, implementers, and end-users. The framework gives the means to make sure that the steps of the IP are performed as intended and to evaluate products and outcomes relevant to using the IP. The evaluation framework informs a system for data collection, with questions mapped to the process, product, and outcome elements.

The IP presents unique challenges to crafting an evaluation. The IP takes a user-centered design approach in bringing multiple disciplines and end-users together to produce written and computable guidelines. During the IP process, there may be jargon used by specialists that may be unfamiliar across disciplines. But the value of the IP is that it makes these interactions more transparent and focuses on the need for regular, planned, cross-disciplinary communication to promote understanding and application of the tool to achieve the desired outcomes.

The evaluation framework, in turn, is designed to appraise this multidisciplinary approach to guideline development and implementation. The framework integrates an assessment of interdisciplinary engagement, standards to assess quality of the guideline and its derivative products, outcome evaluation measures, and user-centered design principles in a single evaluation framework. To our knowledge, this type of evaluation framework has not previously been developed.

Applying the evaluation tool to a case study helped to examine what occurred through all phases and activities and revealed gaps in fully implementing the process. Had the tool been completed at the time of the C19 DGWG effort, as opposed to after the data were collected to support the case study, it would have been much easier to gather process evaluation data along the way versus a post hoc analysis. For product evaluation, the guideline derivatives were informally tested with end-users. Having evaluation prompts for capturing product evaluation data would have been helpful. Having outcome measures already identified also would have been helpful for putting in place ways to measure impact, even simple ones such as capturing downloads to determine the reach of the guidance during dissemination.

There are several potential challenges and limitations in applying this evaluation framework. Operationalizing the evaluation within the context of the IP requires efforts not typically associated with traditional guideline development. For example, guideline developers and implementers may not have experience conducting this type of evaluation. However, the framework provided can assist in training staff to learn how to conduct evaluation activities. Secondly, collecting data needed for the evaluation takes time, expertise, and resources. Data is collected and derived from multiple sources, primarily due to the multiphase, multidisciplinary design of the IP. To minimize this challenge, the evaluation team must be appropriately staffed, with a leader, and a project plan identifying specific roles and responsibilities delineated among participating organizations and stakeholders.

Engaging evaluation expertise, when available, can also assist in making decisions about what data should be collected to meet the needs of the various experts involved in the guideline development or implementation process, and end-users. Because the IP evaluation focuses on evaluating process fidelity, providing a feedback loop, evaluators should have sufficient authority to bring issues to light, with time built in during execution of the IP for the overall

team to make changes in response to evaluation findings. A person with evaluation experience (ie, work experience, education or training conducting evaluations) can ensure important questions are asked.

The extent to which the IP realizes the goal of more efficiently translating new evidence to clinical practice will require more study, making evaluation efforts critical. The IP evaluation framework tool can be used to track how the IP is developed and implemented. It will be particularly interesting to see how various groups begin to implement the IP, the benefits achieved, and the challenges met. One anticipated challenge, the transition of groups currently focused on the development of written guidelines in adding informatics disciplines to their work, may be better managed with the help of ongoing evaluation efforts. Assisting guideline developers in acquiring this capability is an opportunity for governmental entities, professional societies, and others to Improve collaboration and accelerate the anticipated benefits to developing and deploying computable guidelines using the IP. Documenting lessons learned using the IP evaluation tool and sharing them among those who choose to implement the IP can be invaluable to others who wish to develop written and computable guidelines and implement in clinical practice or workflows.

Conclusion

Evaluation is an integral component of the IP. Undertaking the IP within an evaluation framework when the guideline is in development can ensure that processes are carried out as intended or otherwise identify and correct problems in real time. Whether carried out during guideline development or afterward, evaluation grants a means to determine whether products and outcomes were achieved as intended and end-users' needs are met. The evaluation framework described in this manuscript is a flexible tool, adaptable to the needs of users that support use of the IP in developing and implementing timely, high-quality, and highly effective written and computable guidelines.

Conflicts of Interest

The authors have no conflicts of interest to disclose.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention and the Agency for Toxic Substances and Disease Registry.

I.M.L (Centers for Disease Control and Prevention (CDC)) M.M.B (Holistic Health Solutions, LLC), share the role of cosenior author.

References

1. Matson-Koffman DM, Robinson SJ, Jakhmola P, et al. An Integrated Process for Co-Developing and Implementing Written and Computable Clinical Practice Guidelines. *Am J Med Qual*. 2023;38(S2):S12-S34. DOI: 10.1097/JMQ.0000000000000137.
2. Dufour JC, Bouvenot J, Ambrosi P, et al. Textual guidelines versus computable guidelines: a comparative study in the framework of the PRESGUID project in order to appreciate the impact of guideline format on physician compliance. *AMIA Annu Symp Proc*. 2006;2006:219–223.
3. Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E. *Clinical Practice Guidelines We Can Trust*, 2011. https://www.ncbi.nlm.nih.gov/pubmed/24983061
4. Melnyk BM. The Current Research to Evidence-Based Practice Time Gap Is Now 15 Instead of 17 Years: Urgent Action Is Needed. *Worldviews Evid Based Nurs*. 2021;18:318–319.
5. Supplementary tables. An Evaluation Framework for a Novel Process to Codevelop Written and Computable Guidelines. Centers for Disease Control and Prevention. Accessed August 9, 2023. https://stacks.cdc.gov/view/cdc/131007.
6. HL7®. *FHIR® Clinical Guidelines: Agile CPG Development Approach*, 2022. Accessed May 27, 2022. http://hl7.org/fhir/uv/cpg/documentation-approach-05-06-agile-cpg-development-approach.html
7. Michaels M. Adapting Clinical Guidelines for the Digital Age: Summary of a Holistic and Multidisciplinary Approach. *Am J Med Qual*. 2023;38(S2):S3-S11. DOI: 10.1097/JMQ.0000000000000138.
8. Peleg M. Computer-interpretable clinical guidelines: a methodological review. *J Biomed Inform*. 2013;46:744–763.
9. Guyatt GH, Oxman AD, Vist GE, et al; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924–926.
10. USPSTF. *The guide to clinical preventive services: methods and processes. US Preventive Services Task Force*, 2023. Accessed May 22, 2023. https://www.uspreventiveservicestaskforce.org/uspstf/about-uspstf/methods-and-processes
11. NICE. *The guidance manual. National Institute of Clinical Excellence*, 2023. Accessed May 22, 2023. https://www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-pdf-2007970804933
12. HL7®. *FHIR® Clinical Guidelines Examples (FHIR® Clinical Guidelines (v1.0.0) (STU1))*, 2022. Accessed

April 18, 2022. http://hl7.org/fhir/uv/cpg/examples.html
13. Boxwala AA, Rocha BH, Maviglia S, et al. A multilayered framework for disseminating knowledge for computer-based decision support. *J Am Med Inform Assoc*. 2011;18:i132–i139.
14. Consortium TANS. *Appraisal of guidelines for research & evaluation II*, 2022. Accessed March 25, 2022. http://www.agreetrust.org
15. Dearing JW, Cox JG. Diffusion Of Innovations Theory, Principles, And Practice. *Health Aff (Millwood)*. 2018;37:183–190.
16. Cresswell KM, Worth A, Sheikh A. Actor-network theory and its role in understanding the implementation of information technology developments in healthcare. *BMC Med Inform Decis Mak*. 2010;10:67.
17. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology (Oxford)*. 2009;43:535–554.
18. Yen P-Y, Bakken S. Review of health information technology usability study methodologies. *J Am Med Inform Assoc*. 2012;19:413–422.
19. Gulliksen J, Göransson B, Boivie I, et al. Key principles for user-centred systems design. *Behav Inf Technol*. 2003;22:397–409.
20. Davis FD, Venkatesh V. Toward preprototype user acceptance testing of new information systems: implications for software project management. *IEEE Trans Eng Manage*. 2004;51:31–46.
21. Yusof MM, Kuljis J, Papazafeiropoulou A, et al. An evaluation framework for Health Information Systems: human, organization and technology-fit factors (HOT-fit). *Int J Med Inform (Shannon, Ireland)*. 2008;77:386–398.
22. Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: a systematic review and gaps analysis. *Int J Med Inform (Shannon, Ireland)*. 2017;98:56–64.
23. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. 2009;4:50–50.
24. Prevention CfDCa. Framework for program evaluation in public health. *MMWR Recomm Rep*. 1999;48:1–40.
25. Woodring JV, Kruszon-Moran D, Oster AM, et al. Did CDC's 2006 revised HIV testing recommendations make a difference? Evaluation of HIV testing in the US household population, 2003-2010. *J Acquir Immune Defic Syndr (1999)*. 2014;67:331–340.
26. Corporation TM. *COVID-19 Healthcare Coalition*. Accessed March 1, 2022. https://c19hcc.org/.
27. Middleton B, Burton M, Tignanelli CJ, et al. *Pandemic Wakes The Learning Health System: How We Developed And Continue Refining COVID-19 Treatment Guidelines*. 2021(August 26, 2022). https://www.healthaffairs.org/do/10.1377/forefront.20211019.864563.
28. Physicians ACoE. Clinical Resources. *COV-19 ED Management Tool Now Available, Updated*. Accessed August 26, 2022, https://www.acep.org/corona/COVID-19-alert/covid-19-articles/covid-19-ED-management-tool-now-available.
29. HL7®. *FHIR® Clinical Guidelines*. Accessed August 26, 2022. http://hl7.org/implement/standards/fhir/uv/cpg/.
30. HL7®. *FHIR® Clinical Guidelines. Methodology Example: COVID-19 Severity Classification*. Accessed August 29, 2022. http://hl7.org/fhir/uv/cpg/methodology-case-study.html.