ORIGINAL RESEARCH

A Decision-oriented Approach to Evaluating a Leadership Curriculum in Fellowship

Aanchal Kapoor[1], Arunab H. Mehta[4], Salome Arobelidze[2], and Cecile M. Foshee[3]

[1]Department of Critical Care, Respiratory Institute, [2]Department of Hospital Medicine, and [3]Education Institute, Cleveland Clinic, Cleveland, Ohio; [4]Division of Hospital Medicine, Department of Medicine, University of Cincinnati, Cincinnati Medical Center, Cincinnati, Ohio

ORCID IDs: 0000-0001-5130-2373 (A.K.); 0000-0003-0819-421X (A.H.M.)

ABSTRACT

Background: Graduate Medical Education (GME) emphasizes the role of the annual program evaluation to identify opportunities, create action plans, and track improvements longitudinally. There is a lack of a systematic approach to the evaluation of educational curricula. Comprehensive curriculum evaluation can inform the educators about specific modifications to achieve high standards, desired outcomes, and the anticipated objectives.

Objective: To evaluate a leadership in quality improvement program in a pulmonary/critical care fellowship training program using the context, input, process, product (CIPP) model. The CIPP model, given its focus on evaluating different aspects of a program, provides concrete and targeted feedback to guide improvement decisions.

Methods: Evaluation questions addressing the four focused areas were created, pilot tested, and revised. The questions were framed toward optimization of alignment (e.g., program activities with stated objectives, program goals with theoretical perspective, program curriculum with trainee needs) and gaining information about the efficacy of the program in achieving the desired outcomes. To enhance the validity of the results, we triangulated the data-gathering approach by administering surveys and conducting

(Received in original form January 8, 2023; accepted in final form October 26, 2023)

Author Contributions: A.K., A.H.M., and S.A.: Conception and design of study; acquisition, analysis, and interpretation of data; drafting the work and revising it critically for intellectual content; final approval of version to be published; agreement to be accountable for all aspects of the work. C.M.F.: Design of study, helping revise it for intellectual content, final approval of version to be published, agreement to be accountable for all aspects of the work.

Correspondence and requests for reprints should be addressed to Arunab H. Mehta, M.D., M.Ed., 3908 Edwards Road, Unit A, Cincinnati, OH 45209. E-mail: arunab.mehta003@gmail.com.

This article has a data supplement, which is accessible at the Supplements tab.

ATS Scholar Vol 5, Iss 1, pp 96–108, 2024

DOI: 10.34197/ats-scholar.2023-0003OC

interviews and focus groups by random selection from the eligible participants. Qualitative data were transcribed, coded, and categorized into themes aligning with the four aspects of the CIPP model.

Results: We interviewed 9 participants and conducted three focus groups with 20 participants. The surveys provided vital quantitative information that was cross-verified with the qualitative data; 23 of the 25 (92%) participants completed the survey. The results of qualitative thematic analysis were organized in the CIPP format. The context evaluation of the program revealed that the fellows and faculty were unfamiliar with the guiding principles of the course. The input evaluation highlighted the competing interests that hampered the engagement of the fellows during the evening weekly report-outs. The process evaluation revealed clustering of didactic sessions at the start of the course. The product evaluation stressed the difficulty in completing the quality improvement projects in the allotted timeframe.

Conclusion: Conducting a robust evaluation of an educational curriculum provides insights into gaps in the various stages of the program. Time and resources needed for conducting evaluation by using the CIPP model should be considered.

Keywords:
process evaluation; medical residency evaluation; pulmonary critical care evaluation; medical program feedback

Evaluating educational curricula in Graduate Medical Education (GME) programs has been a challenging and time-consuming process. Evaluation of educational programs is necessary to gain insight into the program progress, functionality, and outcomes (1). A systematically structured program evaluation tool can assess whether the program is fulfilling the required needs, is providing anticipated instructions, is delivering the desired outcomes, and is in alignment with the set goals and objectives (2).

Most programs have historically used quantitative data for curriculum evaluation. A recent meta-analysis revealed that qualitative measures were used in only 12% of programs that were evaluated, with Kirkpatrick levels 1 and 2 being the most common forms of measurement (3). Surveys have been the most common form of evaluation. A few programs have implemented preprogram and post-program quantitative assessments to measure the attitude, skills, and knowledge of the learners (4). Only one evaluation was conducted using written and verbal unstructured assessments after rotation in conjunction with learners' board exam performance (5). Some of the evaluations included follow-ups, ranging from surveys sent out 3 months after program completion to telephone calls conducted up to 2 years after program completion (5, 6). Although most surveys collected quantitative data using a Likert scale, a family medicine residency program used qualitative data in the surveys while evaluating quality and leadership curriculum (7).

Recently, a few studies have demonstrated a more complex model-based evaluation. For example, the logic model was used to design and evaluate a quality improvement (QI) leadership course (8). In 2020, a large-scale theory-based evaluation of leadership in a

QI program was performed by using cross-sectional surveys, telephone-based interviews, and database research of participants (9). Two other models of evaluation, appreciative inquiry and logic model, have been successfully used to frame GME program evaluations (10).

Given the systematic, structural, and multidimensional approach of the context, input, process, product (CIPP) framework, we used it to evaluate a leadership in QI program for trainees in the pulmonary and critical care medicine (PCCM) and critical care medicine (CCM) fellowship programs at the Cleveland Clinic to provide robust feedback for program improvement (11).

This 5-month longitudinal curriculum has been designed principally as a leadership course that used QI as the vehicle for teaching. The stakeholders for this program included the fellows who were tasked with establishing an interprofessional team to work on each QI project, coaches who were faculty members with QI experience to guide the fellows while reinforcing leadership attributes when leading the team, and program developers who were the faculty who assisted in developing the program. There were also members of the interprofessional team (such as Nurse Practitioners and Physician Assistants and respiratory therapists who helped the fellows with their projects). The individual stakeholders did not belong to more than one category. These categories were developed based on the philosophy of the Quality in Leadership Program.

The goal of the program was to develop leadership competencies in training fellows; thus, PCCM and CCM fellows were the primary participants of the program. Coaches with QI backgrounds were selected to facilitate QI team projects.

The projects requiring interprofessional team members were selected for the program to provide an opportunity to the fellows to practice leadership behavior while working with team members from diverse backgrounds, professions, and skill sets.

We evaluated the CIPP of both the leadership and QI aspects of the curriculum using a mixed-method approach. This model of evaluation has not been used elsewhere in the institute, to our knowledge.

METHODS

Definitions

The categories of stakeholders are program developers, coaches, participating fellows, and the interprofessional team. Program developers are the stakeholders who developed the goals and objectives of the Quality in Leadership Program based on the guiding philosophy of the program, designed instructional activities and curriculum, and defined roles and responsibilities of other stakeholders in the program, including coaches, interprofessional teams, and participating fellows. The coaches are the faculty members with training and expertise in QI. The coaches assisted the four teams of fellows in their QI projects and monitored and reinforced application of leadership concepts in team interactions. The participating fellows were second-year PCCM and first-year CCM fellows, who were scheduled to participate in the Quality in Leadership Program. The interprofessional team was made up of team members involved in the QI project from interprofessional teams (e.g., nurses, social workers, nurse managers, respiratory therapists, and clinic coordinators). They were an integral part of the project but were not necessarily attending all the program sessions.

CIPP Evaluation Model

The main purpose of evaluation for this program was to assist in formative decisions to help improve the curriculum. The CIPP model (11), given its focus on evaluating different aspects of a program including context, input, process, and product, provides concrete and targeted feedback and hence guides improving decisions (12).

The context evaluation addresses the rationale for the need of the program and focuses on the extent to which the objectives effectively answered the needs of the participants and the organization. It addresses the question, "What should we do?"

The input evaluation focuses on the adequacy of the resources, infrastructure, schedule, and stakeholders to achieve program objectives. It addresses the question, "How should we do it?"

The process evaluation aims to monitor potential problems and areas of strength and collects data on the procedures that happened. It addresses the question, "Are we doing it as planned?"

The product evaluation assists in measuring and interpreting the achievement of objectives and their impact on the participants and the organization. It addresses the question, "Did the program work?"

Evaluation Process Planning

The evaluation of this leadership and QI program for trainees in PCCM and CCM fellowship programs was conducted from April 2021 to August 2021. The study was considered exempt for review by the institutional review board. We proactively engaged key stakeholders from the planning phase to the implementation process, which encouraged their buy-in and ensured that the evaluation was comprehensive, credible, and reflective of their needs and perspectives. We designed evaluation questions to address each stage of the CIPP model (examples are described in Table 1). Evaluation questions serve as a guide to create tailored questionnaires for different data collection tools (e.g., surveys, interviews, and focus groups). Evaluation questions are essential to the use of the CIPP model and the development of subsequent aspects of the evaluation, such as survey questions. These questions were sent to a few experts in the field of education for pilot testing, and the feedback obtained was implemented.

To analyze common themes, understand the context behind unique and dissenting perspectives of different stakeholders, and propose a theory of change, we complemented the quantitative analysis with qualitative data gathering. To enhance the validity, we triangulated our data by obtaining information from multiple sources using multiple methods to address the same questions. Depending on the evaluation objectives, we identified different stakeholders with relevant insights or experiences. Each participant group received a customized data collection method and questionnaire that addressed their specific roles, perspectives, and experiences with the program. The data collection methods used for the evaluation are listed in Table 1.

The cost of evaluation included minor fees for the use of coding software. The three evaluators completed the evaluation process in 3 months and contributed 8 h/wk. We used various methods for collecting evaluation data to get the richest and most diverse information. The method of data collection for each

Table 1. Examples of evaluation questions, data collection methods, and resources

CIPP	CIPP Evaluation Questions	How Will You Measure/What Instruments (Data Collection Tools)	Linked to What Goals/Objectives	Data Collection Resources
C	Are the objectives of the program in alignment with the needs of the fellows?	Surveys/questionnaires to program developers, coaches, and fellows Interviews with program developers and coaches Focus groups with fellows	To equip the fellow trainees of the critical care/pulmonary and critical care departments with the values and skills of leadership and To build a robust quality improvement platform for the fellow trainees to learn the processes involved in quality improvement initiatives and practice leadership behaviors	Fellows, program developers, coaches Getting individualized perspectives on the needs of fellows
I	Are the program activities, content, and material consistent with the fellows' needs, goals, and objectives?	Literature review Focus groups with fellows Interviews with program developers, executors, coaches, experts	Fellows will demonstrate collaboration, teamwork, and stakeholder management skills during quality improvement projects and will be assessed by coaches, 360-degree assessments immediately after completion of the course and 6 months later	Fellows, program developers, executors, coaches, experts For formative assessment of activities and materials
Pc	Are the teaching strategies used by faculty in alignment with program objectives?	Interviews with program developers and coaches	Fellows enrolled in the program will demonstrate leadership behaviors as assessed by the project coaches/mentors during their interactions with the interprofessional teams and during their weekly research in progress report-out sessions	Program developers and coaches To get perspectives on the selection of teaching strategies
Pd	What kinds of evidence will be acceptable by the different stakeholders to define the success of the program?	Interviews with program developers, coaches, program directors, policymakers, expert leaders Focus groups with fellows	Fellows will demonstrate collaboration, teamwork, and stakeholder management skills during quality improvement projects and will be assessed by coaches, 360-degree assessments immediately after completion of the course and 6 months later	Program developers, coaches, directors, policymakers, expert leaders, fellows An assessment of the utility of the evidence for different stakeholders would be appropriate

Definition of abbreviations: C = context; CIPP = context, input, process, product; I = input; Pc = process; Pd = product.
Evaluation questions were created, discussed, and revised addressing the four focused areas based on the CIPP framework. Also mentioned are the sources used for the evaluation as well as the tools used to collect these data (surveys, interviews, focus groups). The complete table is available in the data supplement.

evaluation question was analyzed based on the construct of the question, the interest of the stakeholders, and its added value in triangulation. As described in Table 1, we used surveys, interviews, and focus groups to collect the data.

Data Collection

Surveys. The objective of the survey was to measure attitudes, behaviors, and the application of learned skills. The surveys were administered to the participating fellows, program developers, and coaches.

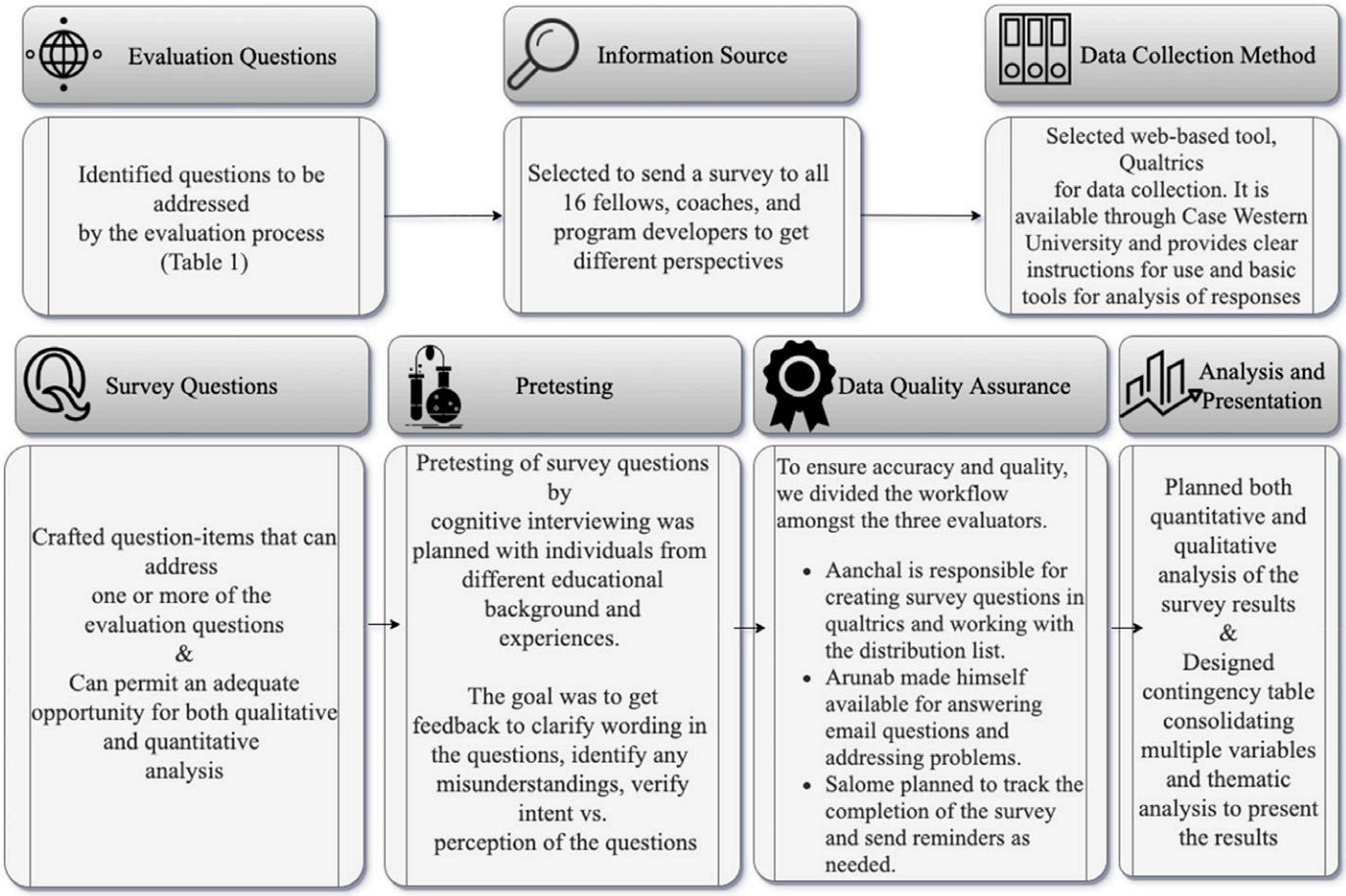

Figure 1. Survey matrix used to help make survey questions.

Before administering the surveys, we conducted a design session with the program developers, coaches, and fellows explaining the intent of the evaluation process, emphasizing the objectives of the program and program theory (13, 14). We ensured the anonymous nature of the survey, explained their right to participate, and guaranteed no repercussions if they decided to opt out.

A survey design matrix was used to summarize the design decisions in a succinct format (Figure 1). We crafted survey question items addressing each evaluation question (Table 1). Each question in the survey was designed to answer one or more evaluation question. We used different item types and open-ended questions to maximize the response and increase the robustness of the survey. While formulating the survey questions, we paid close attention to keeping the survey short, avoiding leading questions, and providing clear instructions for answering the questions.

The survey questions were sent to 10 experts in the educational field for pilot testing. These experts did not participate in the final survey or the program evaluation. They were provided with the overview of the program, evaluation plan, information about the program recipients (fellows), their prior knowledge, and the objectives of the survey. We performed cognitive interviews of the experts, used the "think-aloud" approach, and modified

the questions based on feedback (15). The survey questions were then distributed electronically. The fellows, program developers, and coaches received different versions of the survey.

Interviews. Qualitative interviews are performed to understand the perspectives, attitudes, behaviors, and experiences of others (12). They allow for clarification and probing while permitting exploration and discovery. Interviews are especially useful when the nature of the information is ambiguous and when there is a need for greater depth than what can be obtained through structured surveys.

Thirty-minute interview sessions were conducted, in person and virtually, in a semistructured manner with randomly selected program developers, coaches, and fellows.

The interview questions were formulated to ensure interrater consistency among the three evaluators to direct the interviews while trying to maintain a casual conversation between the interviewer and the interviewee. The descriptive interview question items were based on the evaluation questions that were developed using the CIPP framework.

Focus Groups. Focus groups are qualitative data collection methods that involve organized discussion with a selected group of individuals to gain information about their collective views and provide a rich understanding of participants' experiences and beliefs (16). The aim of the focus groups was to provide an interactive opportunity to participants to describe their reactions to the proposed new program and share the changes they might recommend and to identify any views that can facilitate or hinder the success of the program.

Three focus groups (two focus groups for fellows and one focus group for interdisciplinary teams) were used to gain participants' perspectives and to discover program outcomes, such as how participants applied the knowledge they gained, what barriers they faced, and what changes they would recommend for the future.

Data Analysis

The survey provided vital quantitative information that was cross-verified with the qualitative data obtained from the other methods.

Qualitative data were gathered from the transcribed interviews, focus groups, and narrative items on surveys conducted with the fellows, program developers, coaches, and content experts. These data were transcribed and coded using web-based applications. We created *in vivo* codes by extracting keywords from the excerpts that captured important thoughts or concepts. We then developed a coding scheme and coded the data independently using conventional content analysis. We collectively reviewed, modified, and agreed on the codes. Applying a thematic, inductive approach, we organized the coded data into categories. We organized the categories, performed a thematic analysis, and structured the results in the CIPP format with suggestions for improvement based on the analysis.

RESULTS

The surveys were sent to three program developers, six coaches, and all 16 fellows (8 PCCM and 8 CCM fellows). Fourteen fellows, six coaches, and three program developers completed the survey (23 completed out of 25 sent).

Among program developers, there was broad agreement that the teaching

Table 2. Contingency table consolidating fellows' responses

CIPP Category	Survey Question	Responses (%)				
		Strongly Agree	Somewhat Agree	Neither Agree nor Disagree	Somewhat Disagree	Strongly Disagree
Context	The course is helping you understand important, relevant, and useful leadership concepts.	21.42	42.86	7.14	21.42	7.14
	The course is helping you identify gaps in your leadership skills.	28.57	42.86	7.14	14.29	7.14
Process	You received appropriate guidance from the mentor/coach to lead the project in a meaningful way.	57.14	21.42	14.29	7.14	0

Definition of abbreviation: CIPP = context, input, process, product.

methods used were ideal in helping the learners achieve their learning objectives (100% agreement) and the leadership concepts taught to the learners were helpful to them (all three agreed). The leadership skills that the fellows used most during this course were the team building (nine votes), followed by time management (six votes) and delegation (five votes). Among the coaches, there was consensus that the teaching methods used were ideal in helping the learners achieve their learning objectives (five out of six coaches agreed). A large majority also believed that the fellows were very engaged in the weekly mentorship sessions (five out of six agreed). Most fellows believed that they were achieving appropriate guidance from the coaches during the course (11 out of 14 agreed).

We also found a difference in the responses from PCCM and CCM fellows related to the utility of the course in helping them learn about leadership. We suspect the difference in response is attributed to the post graduate year level of training within the two fellowship program trainees. PCCM fellows participated in this course in the second year of their training; they had experience working with interprofessional teams and got acclimated to the new program teaching them about the foundation of leadership and QI concepts. CCM fellows participated in the course in the first year of their training. They were busy getting themselves familiar with the intensive care unit (ICU) settings, protocols, and clinical work; were likely slower to learn leadership and QI concepts; and felt strained that early in their training.

Table 2 describes the fellows' responses in percentage to each question on a Likert scale (strongly agree to strongly disagree). Figure 2 differentiates the response between the second-year PCCM fellows and first-year CCM fellows.

After transcribing and coding the information obtained from the narrative replies in the surveys, interviews, and focus groups, we categorized the themes and corresponding suggestions in the CIPP format. A total of nine interviews and three focus groups were conducted in addition to 23 survey responses received. Table 3 describes examples of the CIPP model evaluation approach linked to the target areas of interest, with the interpretation of the findings and suggestions for program improvement.

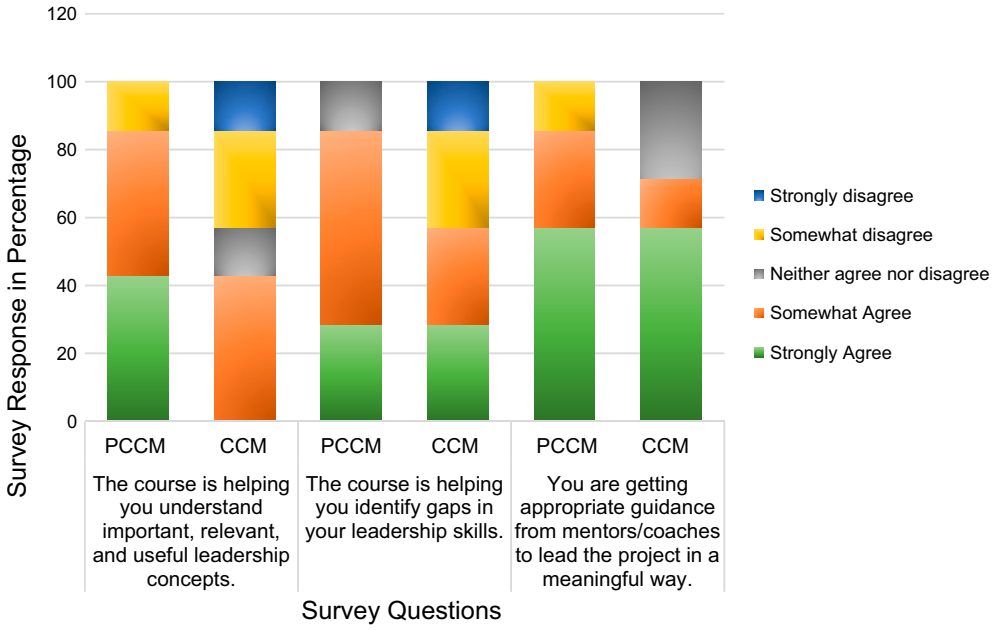

Figure 2. Comparison of responses between pulmonary/critical care medicine (PCCM) fellows and critical care medicine (CCM) fellows in the program.

Brief Summary and Suggested Changes

The context evaluation of the program revealed:

1. Neither the fellows nor the faculty members were aware of the guiding principles of the course. We suggested conducting an orientation before the start of the program.
2. The fellows had difficulty identifying gaps in their leadership skills and believed the program was skewed toward QI. We suggested having a preread on leadership concepts and balancing the amount of leadership and QI teaching even at the end of the program.
3. There were concerns about some fellows being more engaged than others, and leadership opportunities were harder to come by when the team had four fellows. We suggested making the program, or part of it, optional for interested fellows.

The input evaluation of the program highlighted:

1. The competing interests that hampered the engagement of the fellows during the evening weekly report-outs. The concern about a significant time commitment over 5 months was also raised. We recommended prolonging the course and spacing out the reporting sessions based on fellow availability.
2. There was an opportunity identified by program developers to increase interaction between fellows and the interprofessional team and coaches. We suggested creating expectations of the roles of interprofessional teams and coaches before the start of the program.

The process evaluation revealed:

1. Clustering of didactic sessions at the start of the course. We suggested spacing the didactic sessions throughout the course to provide more time in the beginning to start QI projects.
2. There were concerns about additional workload by the first-year CCM fellows while getting acquainted with the ICU environment. We suggested either prolonging the timing of the course or pairing a first-year CCM fellow with a second-year PCCM fellow to provide more support.

The product evaluation stressed:

1. The difficulty in completing the QI projects in the allotted timeframe. We recommended creating multiple tiered goals for each QI project.

Table 3. Example of summative evaluation report using the context, input, process, product framework

CIPP Evaluation Approach	Areas Addressed	Interpretation of Findings	Supporting Comments	Recommendations for the Program
Context (To assess the needs and opportunities within the defined context of program mission)	Leadership	Fellows had difficulty in identifying gaps in leadership skills. Also, they believed the course was more focused on QI.	Fellows: "I think it depends on what everyone wanted, and what each person wanted to get out of it. But it felt like it was still more toward quality improvement."	A preread on leadership competencies and their application might be helpful. In addition, blending the leadership concepts with the QI component rather than keeping the latter part of the program focused on QI will help in redirecting the focus on both components equally.
Input (How to effectively implement the program and determine the resources needed to meet the goals)	Coaches	The coaches provided support and guidance to the teams. It was obvious that mainly the team lead fellow was more in direct contact with the coach compared with the rest of the team.	Fellows: "I don't think I could have asked for a better mentor. If I have another project, it's the same mentor."	Meeting with coaches before the course and together outlining expectations for the entire team would be advised.
Process (How to improve the processes [e.g., teaching activities, program timing, etc.] in the program)	Project report-out	The fellows and coaches broadly believed that the weekly report-out was less valuable, was challenging to prepare, and seemed redundant.	Fellows: "I'm not sure how valuable the weekly recap from each group is. It ends up taking up a lot of time, with an unclear amount of benefit to the other participants. Perhaps instead, each group could provide a recap in a breakout session to a mentor or a faculty."	Presenting a weekly update to respective coaches in the break-out sessions and changing report-out to every other week to a bigger group might be beneficial. One of the groups can present a detailed report-out for 5–7 min every other week, while other teams are giving shorter report-outs for possibly 2 min duration describing the progress from the last session based on a three-slide template.
Product (How to improve the processes to meet program objectives and ensure fellows are using the learned skills)	QI	The fellows faced difficulty completing their QI projects in the time frame of the course.	Coaches: "QI projects have always been a big challenge (because) … most of them never see completion within the given timeframe."	Establishing multiple goals; some of these would be short-term goals that need to be completed at the end of the program, whereas others are long-term goals for which the project could continue beyond the program. Select projects with data already available. Alter the duration of the course.

Definition of abbreviation: CIPP = context, input, process, product; QI = quality improvement.
Summative evaluation report for the program, including suggestions for improvement. A complete table is available in the data supplement.

2. There were no defined metrics to measure the impact of the leadership component of the program, for which we recommended using a 360-degree evaluation for fellows and a follow-up survey regarding their job placement, roles, and positions after the fellowship training.

DISCUSSION

Conducting a robust evaluation of an educational curriculum provided insights into gaps in the various stages of the program implementation.

The CIPP model of evaluation has been used in undergraduate medical education in multiple countries, with variable levels of acceptance of the findings. In 2015, the medical school curriculum at the University of Auckland was revised, with a change in the obstetrics and gynecology clinical attachment from 5 weeks to 4 weeks (17). Applying the CIPP model of evaluation to the revised program enabled identifying the procedures and processes to achieve greater efficiency and maintain the efficacy of the program. Similarly, our evaluation process provided targeted and actionable feedback to be embedded before the subsequent sessions of the program. Although the CIPP model has been used extensively in school districts and state and federal agencies in the United States, it has been used less often for program and/or curriculum evaluation and improvement in GME. The model offers a comprehensive, feasible, and straightforward framework, which makes it a more suitable tool for evaluating diverse educational programs.

This program was developed with the intent of teaching and practicing leadership skills with embedded QI initiatives. The CIPP framework uncovered the difference in perspectives of program developers and other participants in the program. Involving a diverse group of stakeholders in the program evaluation process aided in unfolding the gaps in different aspects of the curriculum.

Various other education evaluation models exist, and Worthen and colleagues in 1997 categorized some commonly used models largely into objective-oriented, consumer-oriented, participant-oriented, and management-oriented evaluation models (18).

The objective-oriented evaluation approach leans on instituting targets before evaluating the program and assessing how successful the program is at meeting those targets. However, the focus is only on evaluations of outcomes such as the end product of student success, which can miss out on the evaluation of the process of education itself, such as teaching and improving (18).

The consumer-oriented evaluation approach stems from the belief that evaluation ought to serve the consumer—that is, the ultimate end user of the program. Therefore, the focus of this evaluation is solely on what the consumers of the educational program want. But this approach can decrease the enthusiasm of educators and curriculum developers because it examines only the consumers' needs (18).

The participant-oriented evaluation approach is one that focuses on the participants in the program, and, hence, the program participants have an interest in the results of the evaluation. It thus empowers all the program participants, even those who have at times been left out of the evaluation. However, subjectivity can limit the utility of the assessment and it is often time and labor intensive (18).

The decision-oriented evaluation approach (such as the CIPP model) provides a comprehensive evaluation of multiple aspects of a program to assist decision makers in taking appropriate alternative steps for program improvement. The main goal of the evaluation is to improve rather than to prove (18).

Strengths

A major strength of the model is that it articulates the stages of a program and

potential information needs at every stage, which helps people, typically managers or policymakers, make decisions. When used for formative purposes, the evaluator might not have to wait for the completion of the program, because the evaluation for context, input, and process can proceed while the program is being designed or is still running its course. Based on the specific needs, either a single component or whole components may be used while evaluating an educational program. Given these advantages, program evaluators should consider choosing this model in preference to other objective-based models (such as Kirkpatrick's model) or consumer-oriented approaches.

Limitations

One of the limitations of the CIPP evaluation model is the time and resources needed for implementing such a model. In terms of the workforce alone, the time and effort needed for the development of questions, data collection, and analysis of said model can be prohibitive. Evaluation using a part of the CIPP model (for example, evaluating just the process) to help focus the evaluation can be used to overcome this limitation.

The CIPP model's effectiveness depends on conducting evaluations at appropriate intervals throughout the program's life cycle. If the evaluations are not conducted in a timely manner, valuable opportunities for program improvement may be missed. This portends another limitation of the model, which can be overcome by scheduling frequent future evaluations.

This approach to evaluation is most effective when the leaders of the program, who are the primary stakeholders and decision makers, actively use the provided information to make decisions regarding program modifications. A limitation of this study could be the single-center nature of this study as well as the limited sample size of the participants (25 participants). More such studies need to be done to enhance the understanding of this method of evaluation.

Future Directions

The CIPP model gave us the novel advantage of providing targeted and actionable feedback to the program developers to be embedded before the subsequent sessions of the program. These changes were implemented for the next academic year with success.

A future process will be to reevaluate the program after the suggested changes. We would hope that given the utility and the information obtained by us using the CIPP model, this would be used extensively in PCCM and CCM educational curricula and other GME programs. Additional studies will add to the knowledge about the utility and implementation of the CIPP model and enhance perceived advantages and disadvantages of this model for GME.

Conclusions

A thorough evaluation of an educational curriculum using the CIPP model yielded valuable insights into gaps across different program stages and diverse stakeholder perspectives. The identified themes and perspectives were effectively communicated to the program developers, who carefully considered them to enhance the program experience for the participants (i.e., fellows in training). Successful implementation of the recommended changes occurred in the following edition of the program.

Author disclosures are available with the text of this article at www.atsjournals.org.

REFERENCES

1. Musick DW. A conceptual model for program evaluation in graduate medical education. *Acad Med* 2006;81:759–765.

2. Posavac EJ. Program evaluation: methods and case studies, 8th edition. New York: Routledge; 2010.

3. Sadowski B, Cantrell S, Barelski A, O'Malley PG, Hartzell JD. Leadership training in graduate medical education: a systematic review. *J Grad Med Educ* 2018;10:134–148.

4. Runnacles J, Moult B, Lachman P. Developing future clinical leaders for quality improvement: experience from a London children's hospital. *BMJ Qual Saf* 2013;22:956–963.

5. Sims KL, Darcy TP. A leadership-management training curriculum for pathology residents. *Am J Clin Pathol* 1997;108:90–95.

6. Karpinski J, Samson L, Moreau K. Residents as leaders: a comprehensive guide to establishing a leadership development program for postgraduate trainees. *MedEdPORTAL* 2015;11:10168.

7. Ellison PA, Hodgson K, Whittaker M-K, Bucknam M, Darling M. Family medicine curriculum in quality improvement, management, and leadership. *MedEdPORTAL* 2012;8:9229.

8. Rajashekara S, Naik AD, Campbell CM, Gregory ME, Rosen T, Engebretson A, *et al.* Using a logic model to design and evaluate a quality improvement leadership course. *Acad Med* 2020;95:1201–1206.

9. Toma M, Blamey A, Mahal D, Gray NM, Allison L, Thakore S, *et al.* Multi-method evaluation of a national clinical fellowship programme to build leadership capacity for quality improvement. *BMJ Open Qual* 2020;9:e000978.

10. Balmer DF, Rama JA, Simpson D. Program evaluation models: evaluating processes and outcomes in graduate medical education. *J Grad Med Educ* 2019;11:99–100.

11. Stufflebeam DL. The CIPP model for program evaluation. In: Madaus GF, Scriven M, Stufflebeam DL, editors. Evaluation models: viewpoints on educational and human services evaluation. Dordrecht: Springer Netherlands; 1983. pp. 117–141.

12. Fitzpatrick JL, Sanders JR, Worthen BR. Program evaluation: alternative approaches and practical guidelines, 4th ed. New York: Pearson; 2010.

13. Glasow PA. Fundamentals of survey research methodology. McLean, Virginia: MITRE, C3 Center; 2005.

14. Ponto J. Understanding and evaluating survey research. *J Adv Pract Oncol* 2015;6:168–171.

15. Ericsson KA, Simon HA. Verbal reports as data. *Psychol Rev* 1980;87:215–251.

16. Gill P, Stewart K, Treasure E, Chadwick B. Methods of data collection in qualitative research: interviews and focus groups. *Br Dent J* 2008;204:291–295.

17. Kool B, Wise MR, Peiris-John R, Sadler L, Mahony F, Wells S. Is the delivery of a quality improvement education programme in obstetrics and gynaecology for final year medical students feasible and still effective in a shortened time frame? BMC Med Educ 2017;17:91.

18. Worthen BR, Sanders JR, Fitzpatrick JL. Educational evaluation: alternative approaches and practical guidelines, 2nd ed. White Plains, NY: Longman; 1997.