Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

WILEY

Process and outcome evaluation of a CBME intervention guided by program theory

Deena M. Hamza PhD[1] | Shelley Ross PhD[2] | Ivy Oandasan MD, MHSc, CCFP, FCFP[3,4]

[1]Postgraduate Medical Education, University of Alberta, Edmonton, Alberta, Canada

[2]Department of Family Medicine, University of Alberta, Edmonton, Alberta, Canada

[3]Department of Education, College of Family Physicians of Canada, Toronto, Ontario, Canada

[4]Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada

Correspondence
Dr Deena M. Hamza, Postgraduate Medical Education, University of Alberta, 2-76 Zeidler Ledcor Center, 8540 112 St NW, Edmonton, Alberta T6G 2L9, Canada.
Email: dh9@ualberta.ca

Abstract

Rationale: Competency-based medical education (CBME) has gained momentum as an improved training model, but literature on outcomes of CBME, including evaluation of implementation processes, is minimal. We present a case for the following: (a) the development of a program theory is essential prior to or in the initial stages of implementation of CBME; (b) the program theory should guide the strategies and methods for evaluation that will answer questions about anticipated and unintended outcomes; and (c) the iterative process of testing assumptions and hypotheses will lead to modifications to the program theory to inform best practices of implementing CBME.

Methods: We use the Triple C Competency-based Curriculum as a worked example to illustrate how process and outcome evaluation, guided by a program theory, can lead to meaningful enhancement of CBME curriculum, assessment, and implementation strategies. Using a mixed methods design, the processes and outcomes of Triple C were explored through surveys, interviews, and historical document review, which captured the experiences of various stakeholders.

Findings: The theory-led program evaluation process was able to identify areas that supported CBME implementation: the value of a strong nondirective national vertical core supporting the transformation in education, program autonomy, and adaptability to pre-existing local context. Areas in need of improvement included the need for ongoing support from College of Family Physicians of Canada (CFPC) and better planning for shifts in program leadership over time.

Conclusions: Deliberately pairing evaluation alongside change is an important activity and, when accomplished, yields valuable information from the experiences of those implementing and experiencing a program. Evaluation and the development of an updated program theory facilitate the introduction of new changes and theories that build on these findings, which also supports the desired goal of contributing toward cumulative science rather than "reinventing the wheel."

KEYWORDS
competency-based education, continuous quality improvement, medical education, outcome evaluation, process evaluation, program evaluation

1 | INTRODUCTION

Health care systems are a product of many factors, and in recent years, medical education has been an area of focus.[1-4] Some evidence suggests current training methods need improvement to support the development of physicians who are able to provide care that meets the needs of complex, diverse, and ever-evolving communities.[1-4] Competency-based medical education (CBME), originally introduced in 1978, has regained momentum over the past decade as an educational approach that can address some of the current issues in health care systems by transforming the way that health care providers are trained.[5,6] While many definitions and interpretations of CBME have been identified in the literature, our discussion of CBME is defined as:

> [...] an approach to preparing physicians for practice that is fundamentally oriented to graduate outcome abilities and organized around competencies derived from an analysis of societal and patient needs. It deemphasizes time-based training and promises greater accountability, flexibility, and learner-centredness.[7]

Although multiple health professions education regulatory bodies around the world have mandated a shift to CBME,[6,8] CBME is not accepted across all stakeholders.[8-11] While the assumptions of CBME are based on established educational theory with supporting evidence,[5,12-15] there have been critical debates in the health professions training literature regarding the potential impact of this education transformation. One of the key debates centres on outcomes: how can we know whether CBME will result in improved outcomes, and can we justify the human and financial costs associated with a transition to CBME when there is no certainty of a better outcome?[16-18]

1.1 | How can we evaluate a complex education intervention, such as CBME?

Implementation is defined as the deliberate introduction of new, or modifying existing, patterns of collective action.[19] These patterns of collective action are institutionally sanctioned, formally defined, consciously planned, and intended to lead to a changed outcome.[19] Implementation of new education models is complex as it involves modifying how people think, act, and organize themselves and others to promote collective action, leading to desired outcomes.[19] Several authors have highlighted the complexity involved in implementing CBME, given the multiple stakeholders and settings that must be involved, along with multiple levels of approvals creating obstacles affecting adoption of CBME.[20-23] Critics have argued that all too often, the complexity involved in implementing change is used as an excuse to justify why CBME cannot be rigorously evaluated.[6,11,24-26] We propose that deliberate evaluation of both the process and outcome of CBME is needed and can be done rigorously, yielding useful information in support of ongoing CBME implementation and improvement.

Criticism in the literature about whether CBME results in "improved" graduates can be answered through use of *outcome evaluation*.[8-11, 16-18] Outcome evaluation examines the progress of the program and the status of accomplishing desired results and answers the questions such as unintended outcomes, return on investment, and changes in knowledge, attitudes, and behaviours.[27,28] While literary conversations about CBME have predominantly focused on outcomes in the form of the individual competence of graduates, understanding best practices in enacting the implementation of CBME is equally as important as measuring its outcomes. In other words, evaluating how CBME is implemented with an understanding of the context within which it is implemented will provide a more a fulsome understanding of why an outcome emerges.

Process evaluation explores social processes and mechanisms during the implementation of an organizational transformation and prospectively draws a bridge to the evaluation of outcomes associated with the change.[29,30] Process evaluation questions can be asked throughout the implementation of CBME with questions such as the following: what are the barriers/facilitators to implementing CBME activities; what has been accomplished; and who is being impacted by the CBME transformation.[27] By understanding factors that influence outcomes, whether positive or negative, more information can be derived for use in the future.

The implementation of CBME is very context dependent. Adapting a CBME intervention that worked in one context may or may not be feasible in another. However, successful uptake of an innovation by others may be more likely if evaluation of both processes and outcomes are planned alongside implementation.[29,31] Without this type of evaluation, information about contextual mechanisms and processes are lacking, and therefore, the risk is that future implementers may fall into the "cargo cult" problem of organizational transformation whereby the process of implementation is reduced to adherence to a "checklist" protocol, but essential mechanisms and processes that supported the outcomes observed in the original context are lacking, resulting in a superficial imitation lacking specificity.[32] Without knowing the "how" and "why" questions related to implementation, implementers may not have access to useful information about the necessary enablers or barriers that impact anticipated outcomes. They also may not know how to carry out implementation in their own contexts if influencing factors are not included within itemized protocols.[29,31,33]

1.2 | Theory-based evaluation

This paper proposes that theory-based evaluation approaches, such as those used in evaluating complex community-based social and health care initiatives,[34-38] can be used for CBME. Theory-based evaluation begins with the development of a program theory (also called a theory of change), which clearly defines a problem that a transformation or intervention is anticipated to address, and how this change (in this

1098 | WILEY Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

HAMZA ET AL.

case the shift to CBME) is anticipated to be successful.[39-42] The program theory defines intended impacts of a change or intervention and then systematically maps factors that contribute to a chain of short- and long-term outcomes that are expected to have impact.[41,43-45] Accompanying the program theory should be a logic model, illustrating the activities involved in CBME implementation including a description of short- and long-term outcomes.[46-49] Contributory factors that influence the process and the outcome of implementation should be identified in the logic model including assumptions/hypotheses of the results of CBME.[50,51]

The program theory acts as a guide for the development of a program evaluation plan that can intentionally explore inputs, processes, outcomes, and impact. The findings from a process and outcome evaluation, as defined in this paper, sets up a cyclical opportunity for ongoing improvement, enhancing the anticipated program theory and its process and outcomes.[52-54] For those involved in CBME, this includes understanding elements influencing the process of evaluating the *fidelity* (ie, does the program as enacted look like the program as conceived?) of implementation, as well as its *integrity* (eg, do workplace-based assessments provide quality feedback that are detailed and actionable?).[29,54-56] Findings from this type of evaluation can enhance our understanding of the best CBME implementation approaches and highlight factors for consideration that may impede or facilitate its success.

1.3 | Worked example: Triple C Competency-based Curriculum (Triple C)

The links between process and outcome evaluation and a clearly defined program theory can best be understood by looking at a worked example. Worked examples describe a problem and the problem-solving approach to arrive at a final solution with a focus on both the outcome and the process.[57-61] Process-oriented worked examples centre on providing an explicit explanation of "how" and "why" certain steps were taken and/or can help explain how they may have contributed to an observed solution.[57-61] Here, we present a worked example from Canadian family medicine residency training of a theory-based program evaluation carried out concurrently with implementation. In addition to illustrating the connections between program theory and evaluation, this worked example shows how uncovering findings to support regular updating of a program theory can facilitate ongoing continuous program improvement. Finally, this worked example can act as a template for individuals looking to design program evaluation for their CBME innovations or other curriculum reforms.

1.4 | Background

The College of Family Physicians of Canada (CFPC) is the accrediting and certifying body for the discipline of family medicine in Canada. In 2010, the CFPC transformed their national family medicine residency

education curriculum and assessment guidelines to align with the core components of CBME.[62] The CFPC's version of CBME is called the Triple C Competency-based Curriculum (Triple C). A detailed description of Triple C can be found in the Triple C Competency-based Curriculum Reports 1 and 2.[63,64]

1.5 | Summary of program evaluation approach

The development of the program theory of Triple C initially focused on identifying how family medicine residency education should be improved by using a social accountability approach to: "address the priority health concerns of the community, region, and/or nation (Canada)."[65] The CFPC consulted with different stakeholders conducted data reviews, and through expert working groups, used a consensus building approach with the academic community to define the family medicine residency education reform to be implemented nationally in Canada from 2010 onwards.[63] CFPC offered Triple C as a guide for use by residency programs as the CFPC had not yet incorporated Triple C into accreditation standards.[63] It was anticipated that adaptation would occur and learnings would be generated that could inform the program theory captured in the program evaluation plan, which would further enhance family medicine's curriculum reform. Figure 1 illustrates the original program theory of Triple C, which is centred on improving the quality of medical education to reach a long-term goal of increased access to and improved quality of care provided by family physicians.

Theory-based evaluations use a logic model to describe key activities and include anticipated short- and long-term outcomes specific to various CBME stakeholders. Figure 2 describes the CFPC's logic model with outcomes hypothesized for family medicine residents, graduates, faculty, the overall discipline, and the CFPC as the accrediting body. (Figure 2)

The logic model (Figure 2) acted as a map to identify opportunities for data collection, which would be incorporated into the evaluation plan that included eight areas of focus evaluating both process and anticipated outcomes (Figure 3). Findings have been used to refine and revise the CFPC's next iterative cycle of curriculum renewal. Learned lessons enhance the original program theory based upon lived experience.

1.6 | Program evaluation: Testing assumptions and revising the original program theory

In Table 1, we illustrate the process of a theory-led program evaluation plan using the original Triple C program theory and subsequent updating based on data collected. The data sources are listed with references shared for those interested in more information. Assumptions from the original program theory and the data source(s) for program evaluation of those assumptions are presented below, along with some examples of evaluation findings and the subsequent refinements to the original program theory that resulted. We do wish to emphasize
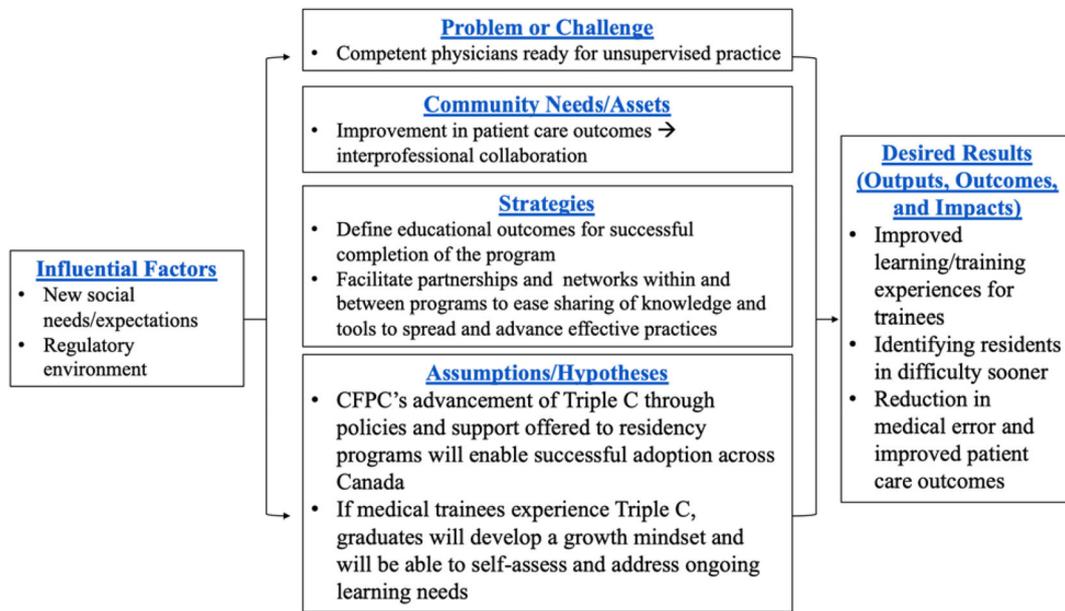
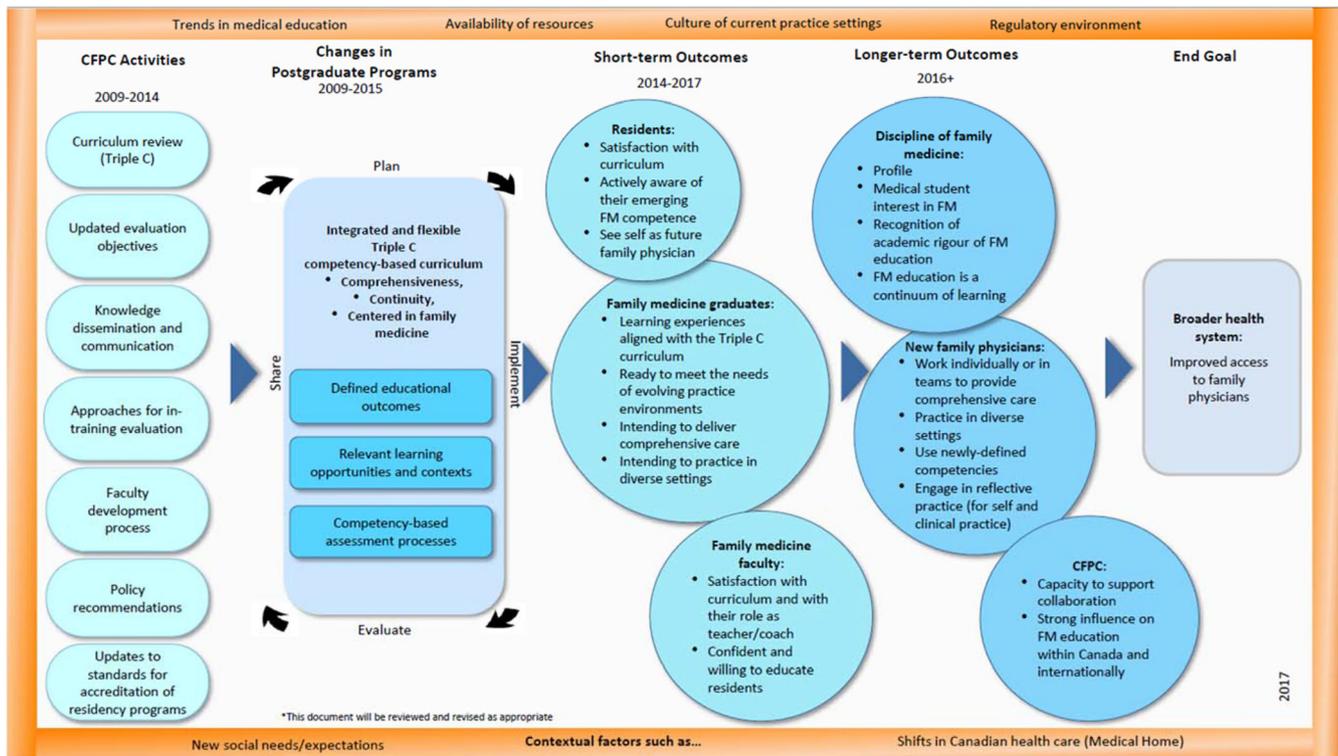**FIGURE 1**   Original Program Theory of Triple C (abridged)

FIGURE 2   Logic model for the Triple C based on strategies, assumptions, and desired results from the program theory

that this is an illustrative example focusing on specific short-term outcomes, and not an exhaustive list of all areas of evaluation included in the full evaluation of the Triple C.

Assumption   The CFPC's advancement of a CBME approach to family medicine residency education through policies (accreditation and certification standards) and change management support

offered to residency programs will enable successful adoption of Triple C across Canada.

Findings from the Residency Program Implementation Profile (RPIP, self-report survey completed by program directors) indicated that all had implemented Triple C across their programs but timing varied.[66] Some early adopters had most core elements of Triple C in

1100 | WILEY — Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research
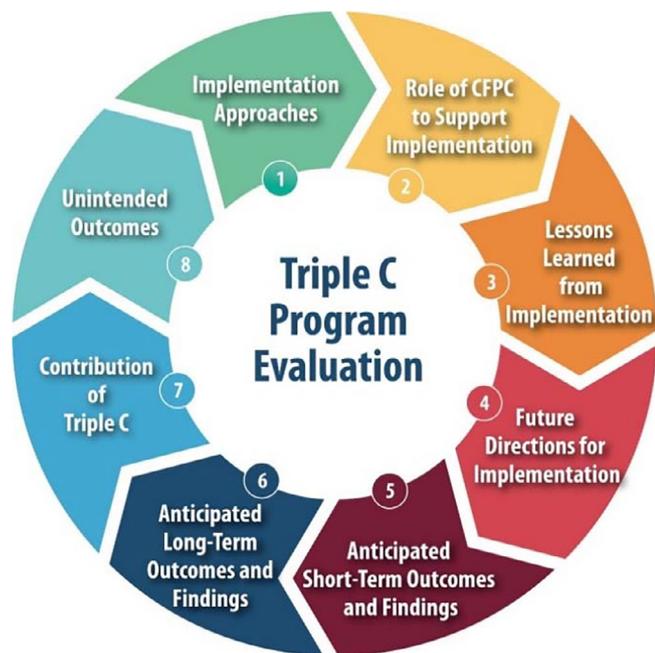
HAMZA ET AL.

FIGURE 3 The Triple C Program Evaluation Plan focusing on implementation, process, and outcome evaluation

place within a year or two of the introduction of this training model, while other programs took several more years to implement Triple C.[66] Findings from the Qualitative Understanding and Evaluation Study of Triple C (QUEST, qualitative study carried out in 2016 examining experiences of Program Directors, Department Chairs, administrative support personnel, postgraduate Deans, and residents through semi-structured interviews and focus groups) uncovered more information related to factors that influenced uptake of Triple C.[67] Early adopters of Triple C reported that the CFPC was helpful in providing support for Triple C through the provision of guides, as well as facilitating opportunities for the sharing of program experiences and lessons learned.[66,68] The co-creation approach taken by the CFPC, collaborating with family medicine program directors and chairs from 2000 to 2010,[63,66,69] also helped increase uptake. While early adopters felt that the CFPC provided adequate support, late adopters, particularly with new program leaders who were not involved in the co-creation of Triple C, felt that the CFPC should have provided more directive and on-going support.[66,68]

1.7 | Updating the original program theory

In Assumption 1, it was hypothesized that CFPC policies and change management support would facilitate Triple C uptake. However, the anticipation of accreditation changes even before the standards were released in 2017 was enough to enable change. The findings suggest that stakeholder involvement in the development of Triple C, especially program leadership, helped to facilitate implementation. The CFPC's guides and early sharing across programs were useful, but later adopters who were not as involved felt the CFPC could have

done more. Effective communication and longitudinal change management strategies to support leadership shifts are a key learning for future use.

Assumption Uptake of Triple C by residency programs will vary depending upon external factors (eg, provincial policies related to family medicine education/practice and medical education culture) and internal factors (eg, faculty engagement, learner demographics, and medical school/residency leadership and infrastructure support).

Findings from the RPIP indicated that the timing of Triple C implementation varied greatly across the 17 university-based programs situated across 10 provinces in Canada and across teaching sites.[66] Although eventual uptake was identified, the time variability was a point of interest. Findings from the Family Medicine Longitudinal Survey (FMLS 2010-2017, self-report survey administered to family medicine residents at entry to and at exit from residency across 16 family medicine programs in Canada) identified that the majority of residents had experienced comprehensive and continuity of care experiences centred in family medicine as well as competency-based programmatic assessment elements specific to Triple C.[66,70] The QUEST study reported that programs with supportive administrative infrastructures, effective communication networks, protected resources (eg, human and financial capital and time), and who had engaged in consultative approaches with stakeholders (eg, postgraduate deans and ministries of health) had smoother Triple C transitions.[66,68] Those who were simultaneously undergoing accreditation reviews and/or expansion of training sites were able to include changes needed to transition to Triple C, capitalizing on pre-existing funding and leadership support.[66,68] Certain factors challenged implementation, such as differing interpretations or definitions of the core components of Triple C.[66,68] Consistent shared mental models amongst all those involved in advancing Triple C were developed through ongoing communication and collaboration between stakeholders such as administrators, program advisors, curriculum designers, and postgraduate deans. These shared mental models helped to drive and support implementation[66,68] and facilitated uptake. Additionally, faculty development for all preceptors coaching family medicine residents was identified unilaterally as an area requiring attention.[66,68]

1.8 | Updating the original program theory

Findings from the RPIP and from the FMLS identified that programs had implemented Triple C. The second assumption noted that both external and internal factors would influence the uptake of Triple C but it did not offer many specifics. Although all programs implemented Triple C, the variability related to time was further understood from the QUEST study. This has prompted the CFPC to consider how best to understand how and if provincial governments and external policies influence residency education and residency

TABLE 1 Program evaluation process showing assumptions from the original program theory, data collection methods, and revisions of the program theory based on findings specific to short-term outcomes

Original Program Theory of Triple C	Data Sources	Updated Program Theory of Triple C based on findings
Assumption 1: the CFPC's advancement of a CBME approach to family medicine residency education through policies (accreditation and certification standards) and support offered to residency programs will enable successful adoption of Triple C across Canada	Residency Program Implementation Profile (RPIP)[a] Qualitative Understanding and Evaluation Study of Triple C (QUEST) Study[b]	• The advancement of Triple C benefited from a nondirective vertical core approach which encouraged uptake from early adopters even before accreditation standards were implemented specifically for Triple C. • Ongoing implementation support by the CFPC was felt to be needed for adopters at later stages. • Collaborative co-creation with stakeholders supports adoption. • Effective communication with all program leaders is imperative.
Assumption 2: uptake of Triple C by residency programs will vary depending upon external factors (eg, provincial policies related to family medicine education/practice and medical education culture) and internal factors (eg, faculty engagement, learner demographics, medical school/residency leadership, and infrastructure support)	Residency Program Implementation Profile (RPIP) Qualitative Understanding and Evaluation Study of Triple C (QUEST) Study Family Medicine Longitudinal Survey (FMLS)[c]	• Flexibility of strategies to implement core features of Triple C increased program autonomy and ownership, and this increased adoption. • Collective sense of accountability to learners and patients supported timely implementation • Protected resources for Triple C reduces strain of implementation processes. • Sharing processes and successful strategies increases efficient and effective use of resources and motivation for change. • Differing interpretations of concepts in Triple C challenges implementation.
Assumption 3: if family medicine trainees experience Triple C, graduates will choose to practice comprehensive family medicine, will choose to work in diverse communities that may be traditionally underserved, and will be able to self-assess and address ongoing learning needs.	Family Medicine Longitudinal Survey (FMLS)[c] p>/p>Pre-Triple C National Physician Survey (NPS; 2010)[d]	• Ongoing evaluation of processes and outcomes uncovered areas requiring action, such as the need for increased learning experiences in certain clinical domains and/or settings, which facilitates dynamic and rapid continuous quality improvement of Triple C. • Program-specific data provided to program directors can be used to undertake local continuous quality improvement.

Note: The full report: Hamza, DM., Oandasan, I., on behalf of the Program Evaluation Advisory Group. Triple C Competency-based Curriculum: Findings Five Years Post-Implementation.

Abbreviations: CBME, competency-based medical education; CFPC, College of Family Physicians of Canada.

[a]Residency Program Implementation Profile (RPIP, 2015): self-report from programs of their triple C implementation.

[b]Qualitative Understanding and Evaluation Study of Triple C (QUEST) Study: qualitative study carried out in 2016 that examined personal experiences of Program Directors, Department Chairs, administrative support personnel, Postgraduate Deans, and residents related to Triple C implementation.

[c]Family Medicine Longitudinal Survey (FMLS, 2010-2017): Self-report survey administered to family medicine residents at entry to program and at graduation from program across 16 family medicine programs in Canada.

[d]Pre-Triple C National Physician Survey (NPS; 2010): Pan-Canadian self-report survey administered to physicians in practice. Questions specific to scope of practice of family physicians were used as pre-Triple C controls.

education reform. Because of the lack of information that was gathered in the program evaluation about this issue, the CFPC is actively exploring how to study these external influences more fully. In terms of the internal factors influencing uptake, it was interesting to hear how programs that took advantage of existing reform processes, accreditation reviews, and new resources to embed Triple C transitioned more easily. This offers new insight on change facilitators.

Assumption If family medicine trainees experience Triple C, then graduates will choose to practice comprehensive family

medicine; will choose to work in diverse communities that may be traditionally underserved; and will be able to self-assess and address ongoing learning needs.

Findings from the FMLS: Residents reported an increase in learning and practice experiences reflecting the Triple C vision after completing their training program.[66] Findings also indicated a few gaps in learning and practice experiences to be addressed by programs and the CFPC.[66]

Findings from the National Physician Survey (NPS 2010, Pan-Canadian self-report survey administered to physicians in practice.

1102 | WILEY— Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

HAMZA ET AL.

Questions specific to scope of practice of family physicians were used as pre-Triple C controls): the NPS illustrated a significant increase in residents' intention to practice comprehensive care in multiple clinical domains and settings by comparing pre- and post-Triple C cohorts.[66,70] Responses from residents also illustrated decreased intentions to practice in certain clinical domains and settings,[66,70] which has prompted the need for additional evaluation and research to understand factors that drive these findings.

1.9 | Updating the original program theory

The demonstration over time that sharing FMLS results back to programs was helpful in implementation of Triple C reinforced the need for program evaluation to run concurrently with implementation. While the initial program evaluation plan was developed to help the CFPC to understand what worked, what did not, and what needed to be changed about Triple C, it was evident early on that the program evaluation data were also immensely beneficial to programs in carrying out local improvement of their curriculum and assessment.

2 | DISCUSSION

Using the Triple C Competency-based Curriculum as a worked example of CBME implementation, we have demonstrated how to develop and update a program theory through theory-based evaluation that looks at both processes and outcomes. This process begins by identifying the challenge or problem, determining the needs and assets of the community, strategies to counter the challenge or problem, and the assumptions that are held that link the strategies to the desired outcomes or results. Influential factors that may impact the problem and outcomes, such as societal needs, are also identified in the development of an original program theory. Once the program theory is developed, the features that should be included in a logic model become evident, such as inputs and activities to support implementation of CBME, process of enacting implementation, and how these initial investments are anticipated to lead to short- and long-term outcomes. We have also illustrated how a program theory and logic model can guide the selection of evaluation methodologies and data collection methods to effectively answer questions about the process and outcomes of implementing new training models in medical education.

While we have presented some examples of the program evaluation process in our worked example, we have also explored how the findings that emerged from our study could be interpreted and used for further improvements for residency program implementation. The CFPC's use of a program evaluation alongside implementation, modelling "collection and action" that uses data on an ongoing basis to inform curriculum implementation and reform, has had multiple benefits to family medicine residency programs in Canada. It is clear that the process and outcomes of a well-designed evaluation can play a significant role in driving ongoing change and ongoing residency

education improvement. Modelling collection and action encouraged engagement in evaluation and research since participants (ie, faculty, teachers, administrators, and residents) had tangible evidence that their efforts supported the growth of Triple C. The outcomes of the program evaluation played a significant role in driving change to improve the future of this discipline.

3 | CONCLUSION

Deliberately pairing evaluation alongside change—such as was done in this CBME education transformation—yields invaluable information from the experiences of those implementing and experiencing a new innovation. Findings "from the field" reinforce the notion that social processes and mechanisms play a significant role in the ways in which a program is experienced and implemented and in how outcomes may differ as a result of these otherwise hidden factors. In addition, findings from program evaluation can support the improvement of current processes, as well as guide future implementation by shedding light on lessons learned and supporting effective and efficient use of resources. Evaluation and the development of an updated program theory also facilitate the introduction of new changes and theories that build on findings, which supports the desired goal of contributing to cumulative science rather than "reinventing the wheel."

Future studies may benefit from ongoing longitudinal evaluation to identify trends in factors that drive, support, and/or challenge change over time and if, how, and why these factors shift over time.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

REFERENCES

1. Donabedian A. Evaluating physician competence. *Bull World Health Organ.* 2000;78:857-860.
2. Horsley T, Lockyer J, Cogo E, Zeiter J, Bursey F, Campbell C. National programmes for validating physician competence and fitness for practice: a scoping review. *BMJ Open.* 2016;6:e010368. https://doi.org/10.1136/bmjopen-2015-010368.
3. Santos P, Alves L, Simoes JA. What distinguishes a competent doctor in medical education? *Int J Med Educ.* 2017;8:270-272. https://doi.org/10.5116/ijme.595f.b2ad.
4. AHRQ. Aligning Medical Education with the Nation's Health Priorities: Innovations in Physician Training in Behavioral and Social Sciences. (2015). <https://www.ahrq.gov/professionals/education/curriculum-tools/population-health/satterfield.html>.
5. McGaghie WC, Miller GE, Sajid AW, Telder TV. Competency-based curriculum development on medical education: an introduction. *Public Health Pap.* 1978;11-91.
6. Holmboe ES et al. A call to action: the controversy of and rationale for competency-based medical education. *Med Teach.* 2017;39:574-581. https://doi.org/10.1080/0142159X.2017.1315067.
7. Frank JR et al. Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Med Teach.* 2010;32:631-637. https://doi.org/10.3109/0142159X.2010.500898.
8. Ferguson PC, Caverzagie KJ, Nousiainen MT, Snell L, Collaborators I. Changing the culture of medical training: an important step toward

the implementation of competency-based medical education. *Med Teach*. 2017;39:599-602.

9. Whitehead, C. Will the triple C curriculum produce better family physicians? No. *Can Fam Physician* 58, 1071, 1073, 1075–1078 (2012).

10. Pimlott, N. Competency-based education. 981–981 (2011).

11. Ross S, Hauer KE, van Melle E. Outcomes are what matter: competency-based medical education gets us to our goal. *MedEdPublish*. 2018;7.

12. Cooke M, Irby DM, Sullivan W, Ludmerer KM. American medical education 100 years after the Flexner report. *N Engl J Med*. 2006;355: 1339-1344. https://doi.org/10.1056/NEJMra055445.

13. Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. *Acad Med*. 2002;77: 361-367.

14. Swanwick, T. & Association for the Study of Medical Education. *Understanding Medical Education: Evidence, Theory, and Practice*. Second Edition

15. Frank JR et al. Competency-based medical education: theory to practice. *Med Teach*. 2010;32:638-645. https://doi.org/10.3109/0142159X.2010.501190.

16. Blinman, T. Competency based medical education is wrong for surgery. (2017).

17. Long DM. Competency-based residency training: the next advance in graduate medical education. *Acad Med*. 2000;75:1178-1183.

18. Albanese M, Mejicano G, Gruppen L. Perspective: competency-based medical education: a defense against the four horsemen of the medical education apocalypse. *Acad Med*. 2008;83:1132-1139.

19. May C, Finch T, Mair F, et al. Understanding the implementation of complex interventions in health care: the normalization process model. *BMC Health Serv Res*. 2007;7:148. https://doi.org/10.1186/1472-6963-7-148.

20. Hawkins, R. E., Welcher, C. M., Holmboe, E.S., Kirk, L. M., Norcini, J. J, Simons, K. B., & Skochelak, S. E. Implementation of competency-based medical education: are we addressing the concerns and challenges? *Med Educ* 49, 1086–1102 (2015).

21. Caverzagie KJ, Nousiainen MT, Ferguson PC, et al. Overarching challenges to the implementation of competency-based medical education. *Med Teach*. 2017;39:588-593.

22. Frank, J. R., Snell, L., Englander, R., Holmboe, E., S., & ICBME Collaborators. Implementing competency-based medical education: moving forward. *Med Teach* 39, 568–573 (2017).

23. Harden RM. Ten key features of the future medical school—not an impossible dream. *Med Teach*. 2018;40:1010-1015.

24. Gruppen L, Frank JR, Lockyer J, et al. Toward a research agenda for competency-based medical education. *Medical Teachers*. 2017;39: 623-630.

25. van Melle E, Gruppen L, Holmboe ES, Flynn L, Oandasan I, Frank JR. Using contribution analysis to evaluate competency-based medical education programs: it's all about rigor in thinking. *Acad Med*. 2017; 92:752-758.

26. Boyd VA, Whitehead CR, Thille P, Ginsburg S, Brydges R, Kuper A. Competency-based medical education: the discourse of infallibility. *Med Educ*. 2018;52:45-57. https://doi.org/10.1111/medu.13467.

27. Boothroyd, R. A. Process and outcome evaluation approaches. (Louis de la Parte Florida Mental Health Institute, University of South Florida, 2018).

28. ACDS. Outcome evaluation: definition and overview. *Measuring the Difference: An outcome evaluation resource for the disability sector* (2009). <https://www.acds.ca/images/webpages/evaluation/MTD_Module_1_Outcome_Evaluation_Definition_and_Overview.pdf>.

29. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350: h1258.

30. Saunders RP, Evans MH, Joshi P. Developing a process-evaluation plan for assessing health promotion program implementation: a how-to guide. *Health Promot Pract*. 2005;62:134-147.

31. Hulscher MEJL, Laurant MGH, Grol RPTM. Process evaluation on quality improvement interventions. *BMJ Qual Saf*. 2003;12:40-46.

32. Feynman RP. Cargo cult science. *Eng Sci*. 1974;37:10-13.

33. Haynes A et al. Protocol for the process evaluation of a complex intervention designed to increase the use of research in health policy and program organisations (the SPIRIT study). *Implement Sci*. 2014;9: 113. https://doi.org/10.1186/s13012-014-0113-0.

34. Donaldson SI, Lipsey MW. *The Handbook of Evaluation: Policies, Programs, and Practices*; 2006:56-75.

35. Judge K, Bauld L. Strong theory, flexible methods: evaluating complex community-based initiatives. *Crit Public Health*. 2001;11:19-38.

36. White H. Theory-based impact evaluation: principles and practice. *J Develop Effect*. 2009;1:271-284.

37. Gaventa J, McGee R. The impact of transparency and accountability initiatives. *Develop Policy Rev*. 2013;31:s3-s28.

38. Davidoff F, Dixon-Woods M, Leviton L, Michie S. Demystifying theory and its use in improvement. *BMJ Qual Saf*. 2015;24:228-238. https://doi.org/10.1136/bmjqs-2014-003627.

39. Davidson EJ. The "Bagging" of theory-based evaluation. *J Multidisciplin Eval*. 2006;3:iii-xiii.

40. Mayne J. Useful theory of change models. *Can J Program Eval*. 2015;30.

41. Rogers PJ. Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*. 2008;14:29-48.

42. Oandasan I, Martin L, McGuire M, Zorzi R. Twelve tips for improvement-oriented evaluation of competency-based medical education. *Med Teach*. 2019;1-6.

43. En ligneDiMaggio PJ, Powell WW. The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *Am Socio Rev*. 1983;48:147-160.

44. Funnell SC, Rogers PJ. *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. Vol 31. John Wiley & Sons; 2011.

45. Astbury B, Leeuw FL. Unpacking black boxes: mechanisms and theory building in evaluation. *Am J Eval*. 2010;31:363-381.

46. WF. Program theory and logic models. (Wilder Foundation, Minnesota, USA, 2009).

47. Oosthuizen C, Louw J. Developing program theory for purveyor programs. *Implement Sci*. 2013;8:23. https://doi.org/10.1186/1748-5908-8-23.

48. Frumkin P, Galaskiewicz J. Institutional isomorphism and public sector organizations. *J Public Admin Res Theor*. 2004;14:283-307.

49. Van Melle E. Using a logic model to assist in the planning, implementation, and evaluation of educational programs. *Acad Med*. 2016;91: 1464.

50. Munter C, Cobb P, Shekell C. The role of program theory in evaluation research: a consideration of the what works clearinghouse standards in the case of mathematics education. *Am J Eval*. 2016;37:7-26.

51. Sharpe G. A review of program theory and theory-based evaluations. *Am Int J Contemp Res*. 2011;1:72-75.

52. HHS. (ed U.S. Department of Health & Human Services) (Office of Adolescent Health, United States, 2018).

53. Kerssens-van Drongelen I. The iterative theory-building process: rationale, principles and evaluation. *Manag Decis*. 2001;39: 503-512.

54. Dixon-Woods M, Bosk CL, Aveling EL, Goeschel CA, Pronovost PJ. Explaining Michigan: developing an ex post theory of a quality improvement program. *Milbank Q*. 2011;89:167-205.

55. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci*. 2007;2:40.

56. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin Psychol Rev*. 1998;18:23-45.

57. Schwonke R, Renkl A, Krieg C, Wittwer J, Alven V, Saldan R. The worked-example effect: not an artefact of lousy control conditions. *Comput Human Behav*. 2009;25:258-266.

1104 | WILEY— Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

HAMZA ET AL.

58. van Gog T, Paas FGWC, van Merrienboer JG. Process-oriented worked examples: improving transfer performance through enhanced understanding. *Instruct Sci*. 2004;32:83-98.

59. van Gog T, Paas F, van Merrienboer JJG. Effects of process-oriented worked examples on troubleshooting transfer performance. *Learn Instruct*. 2006;16:154-164.

60. van Merrienboer JJG, Kirschner PA, Kester L. Taking the load off a learner's mind: instructional design for complex learning. *Education Psychol*. 2003;38:5-13.

61. Brooks CD. *Effects of Process-Oriented and Product-Oriented Worked Examples and Prior Knowledge on Learner Problem Solving and Attitude: A Study in the Domain of Microeconomics*. Florida State University; 2009.

62. van melle E, Frank JR, Holmboe ES, Dagnone D, Stockley D, Sherbino J. A core components framework for evaluating implementation of competency-based medical education programs. *Acad Med*. 2019;94:1002-1009.

63. Oandasan, I., Saucier, D., eds. Triple C Competency-based Curriculum Report—Part 2: advancing implementation. (College of Family Physicians of Canada, Mississauga, ON, 2013).

64. Tannenbaum, D., Kerr, J., Konkin, J., Organek, A., Parsons, E., Saucier, D., Shaw, L., Walsh, A. Triple C Competency-based Curriculum. Report of the Working Group on Postgraduate Curriculum Review - Part 1. (College of Family Physicians of Canada, Mississauga, ON, 2011).

65. Boelen C, Heck JE, WHO. *Defining and measuring the social accountability of medical schools*. Geneva: World Health Organization; 1995.

66. Hamza, D. M., Oandasan, I., on behalf of the Program Evaluation Advisory Group. Triple C Competency-based Curriculum: Findings Five Years Post-Implementation. (Mississauga, ON, 2018).

67. Ellaway RH et al. The impact of a National Competency-Based Medical Education Initiative in family medicine. *Acad Med*. 2018;31:31. https://dx.doi.org/10.1097/ACM.0000000000002387.

68. Ellaway RH, Palacios Mackay M, Lee S, et al. The impact of a National Competency-Based Medical Education Initiative in family medicine. *Acad Med*. 2018;93:1850-1857. https://doi.org/10.1097/ACM.0000000000002387.

69. Inc., C. C. *Evaluation framework for the Triple C Competency-based Curriculum* (College of Family Physicians of Canada, 2012).

70. Zhang PZ, Hamza DM, Ross S, Oandasan I. Exploring change after implementation of family medicine residency curriculum reform. *Fam Med*. 2019;51:331-337. https://doi.org/10.22454/FamMed.2019.427722.