Development, Implementation, and Meta-Evaluation of a National Approach to Programmatic Assessment in Canadian Family Medicine Residency Training

Shelley Ross, PhD, Kathrine Lawrence, MD, Cheri Bethune, MD, MClS,
Theresa van der Goes, MD, Luce Pélissier-Simard, MD, MSc,
Michel Donoff, MD, Thomas Crichton, MD, Thomas Laughlin, MD,
Kiran Dhillon, MD, Martin Potter, MD, MSc, and Karen Schultz, MD

Abstract

The growing international adoption of competency-based medical education has created a desire for descriptions of innovative assessment approaches that generate appropriate and sufficient information to allow for informed, defensible decisions about learner progress. In this article, the authors provide an overview of the development and implementation of the approach to programmatic assessment in postgraduate family medicine training programs in Canada, called Continuous Reflective Assessment for Training (CRAFT). CRAFT is a principles-guided, high-level approach to workplace-based assessment that was intentionally designed to be adaptable to local contexts, including size of program, resources available, and structural enablers and barriers. CRAFT has been implemented in all 17 Canadian family medicine residency programs, with each program taking advantage of the high-level nature of the CRAFT guidelines to create bespoke assessment processes and tools appropriate for their local contexts. Similarities and differences in CRAFT implementation between 5 different family medicine residency training programs, representing both English- and French-language programs from both Western and Eastern Canada, are described. Despite the intentional flexibility of the CRAFT guidelines, notable similarities in assessment processes and procedures across the 5 programs were seen. A meta-evaluation of findings from programs that have published evaluation information supports the value of CRAFT as an effective approach to programmatic assessment. While CRAFT is currently in place in family medicine residency programs in Canada, given its adaptability to different contexts as well as promising evaluation data, the CRAFT approach shows promise for application in other training environments.

Please see the end of this article for information about the authors.

Correspondence should be addressed to Shelley Ross, Department of Family Medicine, University of Alberta, 6-10 University Terrace, Edmonton, AB T6G 1K4, Canada; telephone: (780) 248-1264; email: sross@ualberta.ca; Twitter: @S_RossUofA.

As competency-based medical education (CBME) becomes more widely adopted for health professions education training around the world, there is a growing need for examples of implemented CBME curricula and assessment programs. Specifically, there is a desire for descriptions of innovative approaches to assessment that are generating appropriate and sufficient information to allow programs to make informed, defensible decisions about learner progress. Meeting this need is crucial: to justify the CBME paradigm shift away from traditional approaches to training, it is necessary to show that CBME programs are doing a better job of teaching and assessing the competence of learners compared with previous approaches.[1–4]

In this article, we describe the development and implementation of CBME in family medicine postgraduate training in Canada, with an emphasis on the approach to programmatic assessment. We present a meta-evaluation of findings that support the approach taken in Canadian family medicine training, as well as offer some insights into lessons learned.

Context

In 2010, the College of Family Physicians of Canada (CFPC) introduced CBME to family medicine residency training with the Triple C competency-based curriculum (Triple C).[5] Triple C establishes the curriculum and assessment standards for all family medicine postgraduate programs in Canada and is the basis for family medicine specialty–specific accreditation requirements.[6] Postgraduate (residency) training in Canada is a 24-month program that follows completion of an undergraduate medical degree (e.g., MD). See Supplemental Digital Appendix 1, at http://links.lww.com/ACADMED/B281, for a brief description of both family medicine training in Canada and the Triple C curriculum.

The assessment of competence as guided by Triple C is outlined in the Continuous Reflective Assessment for Training (CRAFT) document,[7] which presents high-level guidelines for competency-based teaching and assessment. In the next sections, we present details of how CRAFT was developed, what it looks like, and how it was implemented.

Development of CRAFT

The CRAFT approach to programmatic assessment is grounded in published evidence from experts in learning and assessment.[7] The developers of CRAFT took a constructivist approach to designing a system of assessment that employs nonstandardized tools alongside defined approaches, with a focus on workplace-based assessment. In developing CRAFT, the working group was guided by the following beliefs: the need to consider the context of the assessment,[8] recognition of the subjective nature of assessors,[9] and the importance of ensuring that multiple observers contribute to assessment because there is value in incorporating a variety of valid viewpoints[10] into workplace-based assessment.[11–13] Along with assessment *of* learning, there is the explicit incorporation of the principles of assessment *for* learning,[14] guided self-assessment,[15] and assessment *as* learning.[16] While assessment *for* learning has become widely recognized as a necessary consideration in designing CBME programs,[11] CRAFT was developed to incorporate assessment *as* learning through the explicit requirement for reflection (the R in CRAFT) as part of developing accurate self-assessment skills in learners. Coaching learners to engage in guided self-assessment[17–19] begins early in postgraduate medical training, and the CRAFT process intentionally engages learners in regular active reflection of both low-stakes (formative) and high-stakes (summative) assessment information.

The goals of CRAFT are twofold:

- Delineate for training programs the requirements of a context-appropriate learning and assessment framework that will ensure that rigorous and defensible assessment leads to graduating competent physicians; and

- Instill in those physicians the skills that they will need to maintain and enhance their competence while in practice, through fostering and promoting self-assessment skills and adaptive approaches to identifying and remedying gaps in knowledge or skills and safely applying that new knowledge to patient care.

The end result is an approach to programmatic assessment in residency training that incorporates multiple assessment elements to allow for informed decision making about each learner's progress in demonstrating competence. The CRAFT model provides a high-level approach to combining learning and assessment activities in an intentional way to design a programmatic assessment system that can be adapted to local contexts, including size of program, resources available, and structural enablers and barriers. The CRAFT framework aligns with the 5 dimensions of programmatic assessment described in the model proposed by van der Vleuten and colleagues in 2012.[13]

What Is CRAFT?

The CRAFT model encompasses both the "what" and the "how" of programmatic assessment in a learner-centered approach. The "what" of CRAFT clearly defines what should be assessed in the 2 years of family medicine residency training, across the CanMEDS-FM roles and the CFPC's essential skill dimensions.[20] The competencies to be assessed are captured in the assessment objectives for family medicine.[21,22] The "how" of CRAFT has 2 components: the structure of the learning and assessment environment and a high-level model of programmatic assessment that delineates the activities and processes that must occur in each residency program. A graphic depiction of the CRAFT model of continuous and recursive programmatic assessment is shown in Figure 1. While all residency programs must design programmatic assessment frameworks that align with the CRAFT model, there is considerable flexibility to allow each program to include specific approaches and tools, assign roles, and delineate processes that are appropriate for their local context, including program structure, location, and resources.

The first aspect of the "how" of CRAFT requires that the learning and assessment environment is structured to ensure continuity of supervision for every learner. Each incoming learner is matched to a continuous advisor, similar to the structure of master's or doctoral graduate programs. As a result, there is continuity of supervision to monitor development and demonstration of competence over time, as the advisor has access to all assessment information about the learner from beginning to end of residency training. Whether the advisor has a direct assessor role or is a nonassessing mentor varies between training programs, according to local context and structure. The continuous advisor acts as a coach for the learner, guiding the learner along the trajectory of development of competence. Learners use their own assessment data collected across multiple contexts and from multiple observers to reflect on their progress, strengths, and gaps. The learner's reflections are discussed with the advisor, who offers suggestions and guidance as the learner engages in this guided self-assessment process and helps the learner to course correct when their interpretation of their own progress does not align with the assessment evidence.

The bespoke aspect of CRAFT (i.e., a high-level national guideline that allows for a programmatic assessment design that is suitably tailored to a specific local program context) is best understood by considering a key resource of CRAFT, the in-training assessment (ITA) template[7] (Table 1). For each step or task in the ITA template, a description of each activity is presented, and the role of the learner is usually described. However, the ITA template is intended to offer high-level guidelines, rather than to be a prescriptive set of rules; as a result, the ITA template specifies program-specific elements to allow for local decisions about what assessment roles will look like. It should be noted, however, that accreditation standards require that any local, program-specific assessment elements must align with the overall CRAFT guidelines and must adhere to the phases of the CRAFT model of programmatic assessment. Further, graduation from residency training should not occur until a learner has successfully demonstrated competence in all required competencies. At that point, the program director attests that the learner has completed residency training and may write the certification examination (a national requirement).

As shown in Figure 1, there are 3 main phases to CRAFT (learn, assess, decide), and each phase has specific elements such as observe, adjust, and adapt. In Table 1, we show how the phases of the CRAFT framework align with the tasks and steps of the ITA template.

One innocuous word encapsulates a key component of CRAFT: the "adapt" element of the decide component (Figure 1). In this context, adapt refers

Table 1
The In-Training Assessment Template[a] Aligned With the Phases and Elements of the CRAFT Model

CRAFT phase and elements	Description	ITA tasks/steps	ITA learner role	ITA faculty/ staff role
Learn				
Observe	Observers watch the learner in the workplace, looking for: competencies demonstrated; populations worked with; learner's knowledge and skills	"*Observe*, document, and provide feedback during daily clinical activities"	"*Seeks opportunities to* be observed, seeks feedback, participates in documentation process"	Program-specific
Document	Representative sample of observations collected as evidence of development/demonstrations of competence	"*Collect and organize documentation* within a framework" (such as a file or electronic portfolio)	"Organizes documented observations according to own needs and program requirements"	Program-specific
Feedback	Observers and learners talk about observation: What went well? What could be improved? How can improvement happen (specifics)?	"Observe, document, and *provide feedback* during daily clinical activities"	"Seeks opportunities to be observed, *seeks feedback,* participates in documentation process"	Program-specific
Assess (for, of, and as learning)				
Review	Learner engages in reflective review of assessment evidence. Learner and advisor meet regularly to discuss assessment evidence and learner progress	"Periodically review and reflect on progress based on all documents available"	"Provides self-assessment, participates in a process of guided self-assessment"	Program-specific
Reflect	Learner and advisor discuss "story" of the learner progress; based on evidence (guided self-assessment)	n/a	n/a	n/a
Adjust	Learner and advisor discuss gaps and strengths; learning plans should be developed early in training and should be revisited regularly	"Adjust and adapt learning activities"	Program-specific	Program-specific
Decide				
Adapt	Learner progress over time guides adjustment +/– adaptation of learning plan; need for extra resources to address gaps or support strengths +/– modification of curriculum and/or assessment are individually determined	n/a	n/a	n/a
Update	Learning plans should be updated regularly, with learner input; updates should be based on assessment evidence and include plans for action	"Update learning plans"	"Actively participates in developing the learning plan"	Program-specific
Report	High-stakes assessment decisions must be documented and reported, and must include supporting evidence	"Report"	Program-specific	Program-specific

Abbreviation: ITA, in-training assessment; n/a, not applicable.
[a]Quotes are taken from the ITA template by Lawrence and colleagues,[7] page 3.

specifically to adjustments to training for a learner, based on assessment evidence as the learner progresses toward competence. Most often, it is assumed to refer to modifications to a learner's training program as a result of gaps identified (i.e., need for a period of training focused on addressing gaps informally, or a formal remediation contract). However, adapting is also used to modify a learner's training program when the learner demonstrates expected competence before the end of a scheduled rotation or even before the end of the 24 months of family medicine residency training. In these cases, the learner's program is modified to allow the learner to focus on getting extended or enhanced clinical experience in areas aligned with their career goals, or to master the clinical competencies beyond the level of competence that the program expects of learners for graduation.

The rationale for adjusting training for highly competent learners rather than allowing for early completion of training is simple: family medicine residency training in Canada is only 24 months long, arguably the shortest postgraduate training program in the world. Becoming a competent physician requires more than demonstrating competence in specific areas at specific intervals; rather, competence must be demonstrated repeatedly and consistently over time and across different contexts. Further, the time that learners spend in postgraduate training allows them to gain experience in a safe environment before they begin unsupervised practice, and contributes to their identity formation as physicians. The flexibility of the adapt component of CRAFT allows for learners to experience individualized training to meet their needs, either through enhanced learning experiences within the minimum time of training (24 months) or through modifications to training to address gaps (including the potential for extension of training beyond 24 months). The one exception for shortening training is the case of transfer learners (learners who transfer out of another specialty residency program and into a family medicine residency program). Transfer learners may undertake a family medicine residency of less than 24 months if they have already demonstrated expected competence in one or more areas
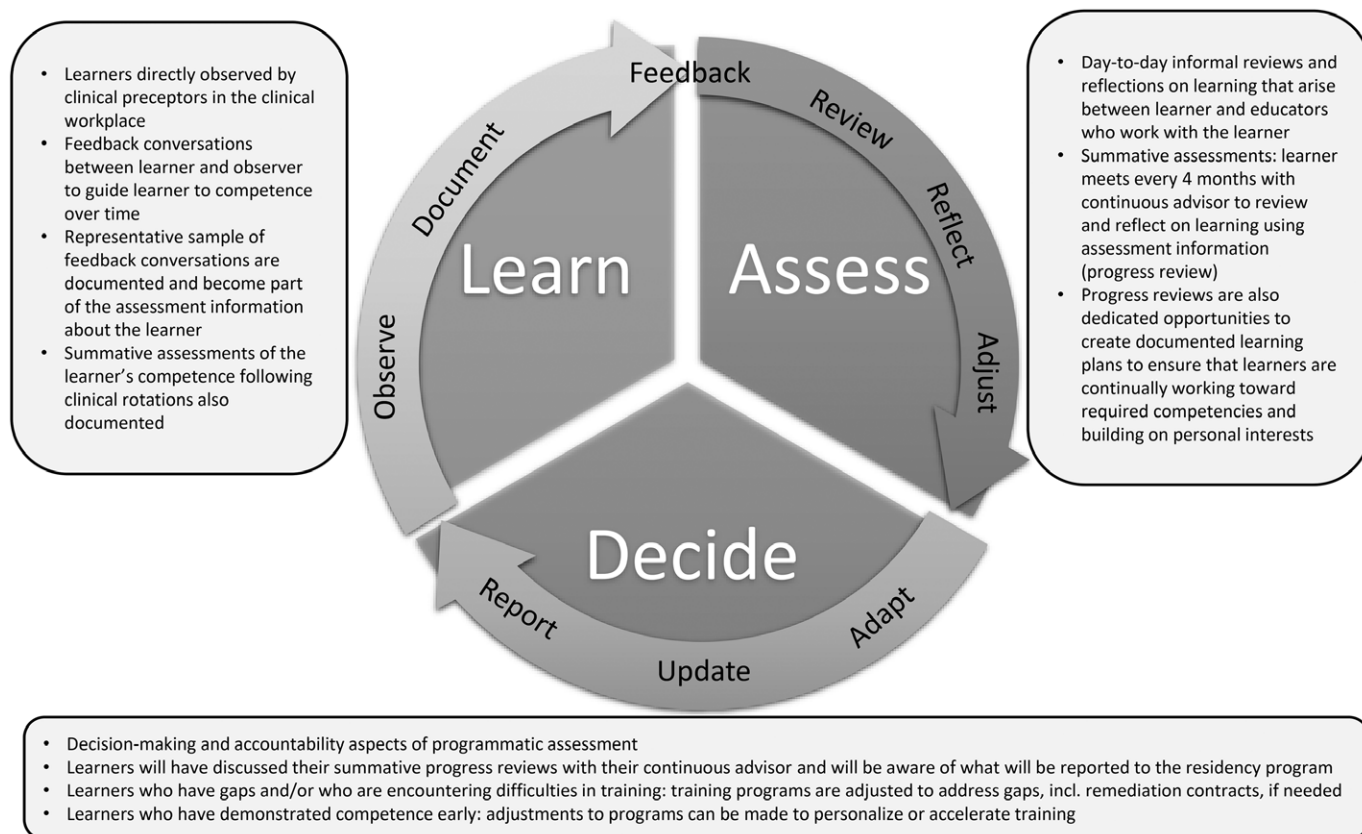
Figure 1 Graphic depiction of the recursive flow of the CRAFT programmatic assessment model. The main phases of the programmatic assessment model are in the middle of the circle, while the activities and tasks associated with each phase are in the outside arrows. The activities and tasks in the outside arrows happen collaboratively between the learner and the teachers who work with the learner. All decisions from the Decide phase are informed by the Assess phase, and feed into the next stage of the Learn phase, in a continuously iterative cycle of programmatic assessment. Abbreviation: CRAFT, Continuous Reflective Assessment for Training.

relevant to family medicine, and the program thus decides that they do not need to complete the associated rotational learning experiences.

Implementation

Implementation of CRAFT occurred between 2010 and 2016. The year of implementation varied across the 17 family medicine residency programs in Canada and depended on local resources and contexts. Each program took advantage of the high-level nature of CRAFT and the ITA guidelines to adapt processes and tools that were appropriate for local contexts.[23–27] An overview of the similarities and differences in CRAFT implementation between 5 different family medicine residency training programs can be seen in Chart 1; detailed information is available in Supplemental Digital Appendix 2, at http://links. lww.com/ACADMED/B281. These 5 programs represent both English- and French-language programs from Western and Eastern Canada.

What is notable about CRAFT implementation across the 5 programs shown is how many similarities there are in processes and procedures. This is despite the intentional flexibility of the CRAFT guidelines, and suggests that allowing programs to adapt processes to suit their local context may result in more uniformity than might be expected. There are 2 major areas where local differences are seen: in the overall structure and nomenclature of local administrative oversight of each program (found in the Description column of Supplemental Digital Appendix 2, at http://links. lww.com/ACADMED/B281), and in the assessment tools and competence frameworks used by each program (found in the first step of the ITA step/tasks).

The remarkable success and speed of implementation of CRAFT in all family medicine training programs across 17 schools and in 2 official languages was likely facilitated by the flexibility inherent in the CRAFT guidelines. This likely reduced barriers to implementation

that can result when program leaders and educators feel that change is being imposed on them, rather than feeling like they have autonomy to make change at their own pace and in a way that works for them and their program.

The other aspect of implementation of CRAFT that contributed to its success was the approach taken by the CFPC. CRAFT was introduced gradually through workshops at conferences and meetings with assessment directors and program directors across the country. These workshops and meetings allowed early adopters to share their successes and failures in venues that allowed for frank discussion of lessons learned. The lessons of early adopters were also openly discussed at the CFPC working groups that contributed to the development of CRAFT, and changes were incorporated into the ongoing development of the CRAFT guidelines. This incremental and iterative approach to implementation of a national guideline meant that by the time CRAFT was formally introduced, many

Chart 1

Overview of Similarities and Differences in CRAFT Implementation for 5 Different Family Medicine Residency Training Programs in Canada, Using the In-Training Assessment Tasks as Illustration[a]

ITA task 1: Observe, document, and provide feedback during daily clinical activities

Common across all programs: Use of FNs (brief narrative capture of feedback shared in the workplace, plus a judgment of performance)

Program-specific variations[b]:

U of A	UBC	U of S	Queen's	Montréal
FN is multi-purpose assessment form	FN is multi-purpose assessment form plus other forms used	Task-specific forms also used	FN is EPA-based	Some FN must include self-reflection

Learner role

Common across all programs:
- Learners encouraged to actively seek feedback
- Learners can create/enter their own FN

Faculty/staff role

Common across all programs:
- All faculty engage in direct observation of learners and share feedback
- All faculty complete FN and other formative assessments

ITA task 2: Collect and organize documentation within a framework

Common across all programs: Online electronic portfolio, or paper/electronic combination. Contains all assessment data (formative and summative) about the learner

Specific organizing framework varies by program:

U of A	UBC	U of S	Queen's	Montréal
Sentinel habits and clinical domains	Essential skill dimensions, domains of care, and priority topics	Information populates periodic review template	Domains of clinical care	CanMEDS-family medicine roles

Learner role

Common across all programs:
- Learners review assessment data; monitor for gaps in assessment data and seek out feedback to address gaps

Faculty/staff role

Common across all programs:
- Continuous advisors monitor information for their continuous learners

ITA task 3: Periodically review and reflect on progress based on all documents available

Common across all programs: Periodic progress review occurs 3 times/year; continuous advisor meets face-to-face with learner

Program-specific variations: Specific process/workflow for periodic review differs by individual program, dependent on local resources (electronic, paper, or a combination of both)

Learner role

Common across all programs:
- Learners prepare for progress review meeting by using assessment data to self-reflect and self-assess

Faculty/staff role

Common across all programs:
- Faculty advisors review learner assessment portfolio to prepare for the meeting

ITA task 4: Adjust and adapt learning activities

Common across all programs:
1. Multiple ways to identify that adaptation of program is needed
2. Continuous advisor and/or program director identify resources and supports to help address learner's needs

Specific processes vary by program

Learner role

Common across all programs:
- Learner is actively involved in discussing concerns and developing a plan to address concerns

Faculty/staff role

Common across all programs:
- Continuous advisor supports learner
- Additional faculty may work with learner (assessing or non-assessing coaching role), as needed

ITA task 5: Update learning plans

Common across all programs:
1. Creation of a learning plan occurs prior to first progress review meeting
2. Review and updating of previous learning plans is part of every subsequent progress review meeting—learning plans are part of the progress review documentation

Learner role

Common across all programs:
- Learner is actively involved in developing initial learning plan, and in reviewing and updating learning plan throughout training

Faculty/staff role

Common across all programs:
- Continuous advisor helps learner through guided self-assessment in creating and revising learning plan

ITA task 6: Report

Reporting processes and procedures are determined by individual programs, and must align with individual university policies

Learner role

Common across all programs:
- Learner is actively engaged in discussion of assessment decision, and signs off that decision was discussed (but can appeal)

Faculty/staff role

Common across all programs:
- Advisor actively engages learner in discussion of summative decision; advisor signs off on decision of progress on the progress review form

Abbreviations: CRAFT, Continuous Reflective Assessment for Training; ITA, in-training assessment; FN, field note; EPA, entrustable professional activity; U of A, University of Alberta; UBC, University of British Columbia; U of S, University of Saskatchewan; Queen's, Queen's University; Montréal, Université de Montréal.

[a]Detailed information for each ITA task, including program-specific variations, can be found in Supplemental Digital Appendix 2, at http://links.lww.com/ACADMED/B281.

[b]Each program is identified by the university with which it is affiliated.

Table 2

Institutional Audit Evaluation Data From 4 Canadian Family Medicine Residency Programs[a]

Evaluation evidence per learner across 2 years of training	Program			
	U of A	UBC	U of S	Queen's
Year CRAFT implemented	2010	2013	2014	2010
Total residents graduated since implementation[b]	~912	~1,119	~322	~650
Mean no. field notes[c]	76	92	72	76
Mean no. other low-stakes assessments	20	27	22	30
Selected examples of other low-stakes assessments	BEAR assignments	Academic integrity module; cultural competence module; video review	Patient and team feedback forms; procedures key features form	Multisource feedback; global health modules; nightmares simulation course
Low-stakes assessments common across all programs[d]	Yes	Yes	Yes	Yes
Mean no. high-stakes (summative) assessments (excluding summative progress reviews)	21	22	14	12
Mean no. periodic progress reviews (mean no. pre-CRAFT)	6 (4)	6 (0)	6 (2)	6 (0)
Mean no. learning plans (mean no. pre-CRAFT)	6 (0)	6 (0)	6 (0)	6 (0)
Approximate % learners/ cohort[e] who need extension of training (nonmedical)	5	Data not available	5	3
Change in numbers of assessments early in training that indicate a concern about a learner	Significant increase[27]	Data not available	Observed (anecdotal) increase	Data not available
Mean no. months to complete residency training	24	24	24	24
Option to individualize training within 24 months of training based on competence demonstrated?	Yes	Yes	Yes	Yes

Abbreviations: CRAFT, Continuous Reflective Assessment for Training; U of A, University of Alberta; UBC, University of British Columbia; U of S, University of Saskatchewan; Queen's, Queen's University; BEAR, Brief Evidence-based Assessment of Research; ALARM, Advances in Labor and Risk Management; ACLS, Advanced Cardiac Life Support.
[a]Availability of data varies by program, due to multiple factors including embargo of data (if the data is being used in a manuscript under review or pending publication), data sharing policies of institutions, and resource capacity within programs to compile specific data.
[b]Core program refers to the 2-year family medicine residency program (see Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/B281).
[c]Based on data from 2018–2019, 2019–2020, and 2020–2021 academic years, except UBC (based on 2019–2020 and 2020–2021 only).
[d]Simulated office oral and short answer management problems (certification examination practice) sessions; practice quality improvement project; ALARM and ACLS courses; resident-as-teacher assessments; narrative reflective practice sessions; ethics modules; scholarly project.
[e]A cohort is defined as the group of residents who begin the residency program as new residents each year.

programs had shared local innovations, and other programs had adapted those innovations for their own contexts. While there were, and continue to be, problems encountered by programs as they fully transform their programs to be competency-based,[23] openly sharing ideas and solutions between residency programs eased the transition significantly.

Meta-Evaluation of CRAFT
Evaluation data for CRAFT come from multiple sources, making a meta-evaluation[28] an appropriate approach. Using a theory-led, outcomes-focused program evaluation framework, our meta-evaluation focused on the assumptions of CRAFT using 2 sources of data: an institutional audit of assessment

information and processes in 4 programs that were early adopters of CRAFT (Table 2), and published evaluation data (Appendix 1). In discussing published evaluation findings, it is important to note that in many publications, authors tend to use the local name for their assessment framework (e.g., Portfolio Assessment Support System[26] and Competency-Based

Achievement System [CBAS][27]) rather than the CRAFT guidelines language, or they refer to the Triple C.

The published evaluation data for the meta-evaluation came from an existing repository of 13 publications (including a white paper and a conference presentation) related to Triple C or CRAFT that the author group created. One author (S.R.) then searched MEDLINE, EMBASE, and PsycInfo with a time filter of 2010–2022 for other publications, using the keywords *CBME, competency-based, assessment, family medicine,* and *Canada.* Inclusion criteria were articles about Triple C and/ or CRAFT in Canadian family medicine residency programs that included evaluation data; exclusion criteria were reviews, program descriptions without evaluation data, and articles about other specialties or undergraduate learners. The same author (S.R.) scanned the titles and abstracts of the 35 results and identified 2 more publications for inclusion that were not in the original repository. The data for the institutional audit were provided by program faculty who were or are directly involved in assessment in each program, and who are also co-authors of this article (S.R., K.L., K.S., T.v.d.G., M.P.). The data were analyzed using qualitative content analysis to summarize the existing evaluation findings and align those findings with 3 key assumptions of CRAFT. This project was deemed exempt from need for ethics review based on the ARECCI screening tool (A pRoject Ethics Community Consensus Initiative).

Assumption 1

Learners are matched to a continuous faculty advisor, giving each learner a dedicated advisor to act as a competency coach. This advisor helps guide the learner to monitor their learning and progress over time, and to support the learner with coaching and guided self-assessment.

Evaluation evidence to support assumption 1 included 4 publications reporting evidence that supported the value of each learner having a continuous advisor to coach and support the learner, including improved skills in reflection and self-assessment,[29,30] significant improvements in addressing identified behavioral and/or learning concerns,[27] and resident self-reports of increased

confidence in identifying personal learning needs.[31] Multiple publications reported associations between the advisor–resident relationship and the identification of residents in difficulty.[23,24,27,31] At a national level, Oandasan and colleagues reported on findings from a longitudinal self-report survey conducted with learners in family medicine training programs in Canada. In comparing responses from learners at the start of training and at the end of training, a significant difference was seen in learners' self-reported ability to identify their own learning needs,[32] and residents' self-reported feelings of preparedness for practice were significantly associated with agreement that the resident had "an identified person guiding my development."[33] Finally, indirect evidence for assumption 1 was provided in a publication in which the authors associated self-reported mastery mindset with the coaching and guided self-assessment experienced by a cohort of residents who graduated from a CBAS residency program.[34]

Assumption 2

Assessment data (especially formative) should come from multiple observers across multiple contexts to give an accurate picture of the strengths and gaps of each learner. As much as possible, assessment data should be based on direct observation in the clinical workplace and should be transparently communicated with the learner through feedback and documentation to facilitate and support learning.

Evidence for assumption 2 addressed the value of assessment data from multiple observers and the increase in direct observation and opportunities for feedback. Ross and colleagues reported significant increases in early detection and remediation of learners in difficulty when comparing data from before and after implementation of CBAS.[27] Other publications also included evaluation evidence for increased identification of residents who were struggling.[23,31,35] This finding has also been anecdotally observed (but not formally evaluated) in other programs across Canada (Table 2). Multiple authors reported evidence for increases in feedback to support learning.[23–26,35–37] Griffiths and colleagues found that regular feedback had become part of the culture of assessment in their program,[29] a finding also reported by Ross and colleagues.[25] In a longitudinal

national survey of graduating residents from 2015 to 2019, over 85% of residents in each cohort self-reported agreement with the statement "I had many informal opportunities for feedback."[32,33] However, it should be noted that in all of the publications cited here except the longitudinal survey,[32,33] a need for more faculty development was also reported.

Assumption 3

The CRAFT guidelines are intentionally written to be high level, rather than specific and concrete. This is so that programs can design and implement programmatic assessment systems that align with the guidelines, but use tools and processes that are specifically appropriate for each individual program's resources and context.

Schultz and colleagues reported on specific facilitating factors that aided the change management needed for successful implementation of their local iteration of CRAFT.[26,31] Two studies that examined data about Triple C implementation included specific data regarding CRAFT. In the evaluation worked example by Hamza and colleagues, the authors identified the high-level guidelines for curriculum and assessment as facilitators of successful implementation, partly by contributing to individual programs' sense of autonomy and ownership.[38] A multi-institution realist study by Ellaway and colleagues also found that the high-level approach rather than proscriptive directions facilitated uptake by allowing programs to adapt and adopt processes and tools for their local contexts.[23] However, Ellaway and colleagues did identify that the wide variation among programs may be problematic.

An overview of specific evaluation data from 4 early-adopter programs (Table 2) shows since implementation of CRAFT, there is consistency across programs in the numbers of progress reviews: now, all programs require a summative progress review meeting and completion of a progress review report every 4 months. Before implementation of CRAFT, progress reviews occurred less frequently in all programs. Further, implementation of CRAFT also resulted in the introduction of learning plans in all programs. In

looking at the amount and type of assessment data collected about each learner over the 2 years of training, some minor variation is seen. This is because the CRAFT guidelines do not dictate specific numbers for different types of assessments. Rather, programs have the ability to adjust numbers of assessments dependent on each learner's trajectory in the development of competence. Learners who consistently demonstrate competence do not necessarily need to be assessed (formatively or summatively) as frequently as do learners who are encountering difficulty. This allowance for learner-centered adjustments to assessment is a result of the emphasis on learning that is embedded in the development of CRAFT.

While the findings from our meta-evaluation support the belief that CRAFT is having a beneficial effect on family medicine residency training, more findings are needed from large-scale studies across multiple programs, as well as long-term outcomes of graduates in practice.

Limitations of the meta-evaluation

A limitation of this meta-evaluation is that while individual programs carry out program evaluation, very few of those program evaluation findings are published. The findings from these projects have been shared among family medicine program directors and assessment directors at the CFPC's Family Medicine Forum conference, including at the assessment directors' meeting held during the Family Medicine Forum, but minimal findings have been disseminated more broadly. A further limitation is that few programs are able to compare outcomes using pre- versus postimplementation data for CRAFT, primarily due to how long CRAFT has been in place (over 10 years in many programs). While some authors have been able to conduct and publish comparison studies (see Ross and colleagues[27]), many programs either do not retain assessment data going back more than a decade, or do not have local capacity or expertise to conduct secondary data analysis of archived assessment data.

CRAFT in Other Contexts

While the CRAFT guidelines are intended to inform design of programmatic assessment in family medicine residency training, the key assumptions of CRAFT are based in educational psychology theory and learning principles, making them transferrable to multiple contexts. Further, CRAFT is not tied to a specific competency framework, nor does it dictate specific tools, so it can be used in any specialty or health care profession and within any educational context. For example, CRAFT could guide programmatic assessment in undergraduate medical education. Students entering medical school would be assigned to a continuous advisor for the duration of medical school. This continuous advisor would form a relationship or educational alliance[39] with the student, and would have access to all assessment data about the student. The continuous advisor and the student would have regularly scheduled meetings to discuss the student's progress, and the advisor would support and guide the student in reflection and guided self-assessment and be responsible for contributing an informed judgment of the student's progress. Assessment information in medical school could expand beyond traditional summative assessments and examinations to include multiple formative assessments based on direct observation that would facilitate feedback conversations, as was seen by Griffiths and colleagues 3 years after their institution's implementation of CRAFT.[29] While this adoption of CRAFT would certainly take work, it would be a transformative way to approach undergraduate medical education.

Similarly, CRAFT could be used in almost any specialty—with the same caveat as above about the work required to introduce this model of programmatic assessment. The easy part would be identifying tools to use for assessment: many specialty programs use entrustable professional activities for assessment, which are already in use in some Canadian family medicine residency programs within the CRAFT model.[35] The most challenging part of implementation would be revising established institutional structures to increase direct observation of learners, and introducing regular formative assessment with accompanying feedback conversations.

Concluding Remarks

CRAFT offers a framework for programs to design learning environments and assessment processes that are intended to support the development of skills needed for lifelong learning—especially self-assessment and self-monitoring of the safe application of new skills and knowledge. Further evaluation data are needed to build upon existing findings.

While CRAFT is currently in place in family medicine residency programs in Canada, this model can be adapted for application in other training environments. The high-level guidelines of CRAFT and the ITA template allow for adaptation to local contexts and resources in settings beyond family medicine residency training in Canada.

S. Ross is professor and director, Research and Innovation, Teaching and Assessment Support Program, Department of Family Medicine, University of Alberta, Edmonton, Alberta, Canada; ORCID: http://orcid.org/0000-0001-9581-3191.

K. Lawrence is associate professor and assessment director and provincial head, Family Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

C. Bethune is professor, Northern Ontario School of Medicine, clinical professor, Memorial University of Newfoundland, Newfoundland, Canada, and clinician educator, College of Family Physicians of Canada; ORCID: http://orcid.org/0000-0002-6230-6262.

T. van der Goes is family physician (retired), medical educator, and director of assessment, University of British Columbia Family Medicine Residency Program, Vancouver, British Columbia, Canada.

L. Pélissier-Simard is associate professor, Department of Family Medicine and Emergency Medicine, and associate academic director, Centre de Développement Professionnel, Faculté de médecine et des sciences de la santé de l'Université de

Sherbrooke, Sherbrooke, Québec, Canada; ORCID: http://orcid.org/0000-0002-9402-1798.

M. Donoff is family physician, professor, and associate chair, Department of Family Medicine, University of Alberta, Edmonton, Alberta, Canada.

T. Crichton is family physician and senior advisor, Postgraduate Medical Education, Northern Ontario School of Medicine, Thunder Bay, Ontario, Canada.

T. Laughlin is an associate professor, Department of Family Medicine, Dalhousie University, Halifax, Nova Scotia, Canada, and clinical associate professor, Discipline of Family Medicine, Memorial University of Newfoundland, Newfoundland, Canada.

K. Dhillon is clinical lecturer, Department of Family Medicine, University of Alberta, Edmonton, Alberta, Canada, and member, Certification Process and Assessment Committee, College of Family Physicians of Canada.

M. Potter is assistant professor, Family Medicine and Emergency Department, Université de Montréal, Montréal, Québec, Canada.

K. Schultz is professor and assessment director, Department of Family Medicine, Queen's University, Kingston, Ontario, Canada, and chair, Certification Process and Assessment Committee, College of Family Physicians of Canada; ORCID: http://orcid.org/0000-0003-0208-3981.

References

1 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.

2 Whitehead C, Kuper A. Faith-based medical education. Adv Health Sci Educ Theory Pract. 2017;22:1–3.

3 Holmboe ES, Sherbino J, Englander R, Snell L, Frank JR; ICBME Collaborators. A call to action: The controversy of and rationale for competency-based medical education. Med Teach. 2017;39:574–581.

4 Dauphinee WD, Boulet JR, Norcini JJ. Considerations that will determine if competency-based assessment is a sustainable innovation. Adv Health Sci Educ. 2019;24:413–421.

5 Oandasan I, Saucier D, eds. Triple C Competency-Based Curriculum Report. Part 2: Advancing Implementation. Mississauga, ON: College of Family Physicians of Canada; 2013. www.cfpc.ca/uploadedFiles/Education/_PDFs/TripleC_Report_pt2.pdf. Accessed May 14, 2022.

6 College of Family Physicians of Canada. Standards of Accreditation for Residency Programs in Family Medicine. Mississauga, ON: College of Family Physicians of Canada; 2018.

7 Lawrence K, van der Goes T, Crichton T, et al. Continuous Reflective Assessment for Training (CRAFT): A national programmatic assessment model for family medicine. Mississauga, ON: College of Family Physicians of Canada; 2018. https://www.cfpc.ca/CFPC/media/Resources/Faculty-Development/CRAFT_ENG_Final_Aug27.pdf. Accessed May 31, 2022.

8 Bates J, Ellaway RH. Mapping the dark matter of context: A conceptual scoping review. Med Educ. 2016;50:807–816.

9 Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the "black box" differently: Assessor cognition from three research perspectives. Med Educ. 2014;48:1055–1068.

10 Govaerts M, van der Vleuten CP. Validity in work-based assessment: Expanding our horizons. Med Educ. 2013;47:1164–1174.

11 Schuwirth L, Ash J. Assessing tomorrow's learners: In competency-based education only a radically different holistic method of assessment will work. Six things we could forget. Med Teach. 2013;35:555–559.

12 van der Vleuten CP. The assessment of professional competence: Developments, research and practical implications. Adv Health Sci Educ. 1996;1:41–67.

13 van der Vleuten CP, Schuwirth LW, Driessen EW, et al. A model for programmatic assessment fit for purpose. Med Teach. 2012;34:205–214.

14 Black P, Wiliam D. Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan. 1998;80:139–148.

15 Sargeant J, Armson H, Chesluk B, et al. The processes and dimensions of informed self-assessment: A conceptual model. Acad Med. 2010;85:1212–1220.

16 Torrance H. Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. Assess Educ. 2007;14:281–294.

17 Duffy FD, Holmboe ES. Self-assessment in lifelong learning and improving performance in practice: Physician know thyself. JAMA. 2006;296:1137–1139.

18 Epstein RM. Assessment in medical education. NEJM. 2007;356:387–396.

19 Sargeant J, Lockyer J, Mann K, et al. Facilitated reflective performance feedback: Developing an evidence-and theory-based model that builds relationship, explores reactions and content, and coaches for performance change (R2C2). Acad Med. 2015;90:1698–1706.

20 Allen T, Brailovsky C, Rainsberry P, et al. Defining competency-based evaluation objectives in family medicine: Dimensions of competence and priority topics for assessment. Canad Fam Phys. 2011;57:e331–e340.

21 Lawrence K, Allen T, Brailovsky C, et al. Defining competency-based evaluation objectives in family medicine: Key-feature approach. Canad Fam Phys. 2011;57:e373–e380.

22 Allen T, Bethune C, Brailovsky C, et al. Defining Competence for the Purposes of Certification by the College of Family Physicians of Canada: The Evaluation Objectives in Family Medicine. Mississauga, ON: College of Family Physicians of Canada; 2010. https://www.cfpc.ca/CFPC/media/Resources/Education/Definition-of-Competence-Complete-Document-with-skills-and-phases.pdf. Accessed June 15, 2022.

23 Ellaway RH, Mackay MP, Lee S, et al. The impact of a national competency-based medical education initiative in family medicine. Acad Med. 2018;93:1850–1857.

24 Ross S, Poth CN, Donoff M, et al. Competency-based achievement system: Using formative feedback to teach and assess family medicine residents' skills. Canad Fam Phys. 2011;79:e323–e330.

25 Ross S, Poth CA, Donoff MG, et al. Involving users in the refinement of the competency-based achievement system: An innovative approach to competency-based assessment. Med Teach. 2012;34:e143–e147.

26 Schultz K, Griffiths J. Implementing competency-based medical education in a postgraduate family medicine residency training program: A stepwise approach, facilitating factors, and processes or steps that would have been helpful. Acad Med. 2016;91:685–689.

27 Ross S, Binczyk NM, Hamza DM, et al. Association of a competency-based assessment system with identification of and support for medical residents in difficulty. JAMA Network Open. 2018;1:e184581.

28 Hedler HC, Gibram N. The contribution of metaevaluation to program evaluation: Proposition of a model. J Multidiscipl Eval. 2009;6:210–213.

29 Griffiths J, Dalgarno N, Schultz K, Han H, van Melle E. Competency-based medical education implementation: Are we transforming the culture of assessment? Med Teach. 2019;41:811–818.

30 Schultz K, McGregor T, Pincock R, Nichols K, Jain S, Pariag J. Discrepancies between preceptor and resident performance assessment: Using an electronic formative assessment tool to improve residents' self-assessment skills. Acad Med. 2021;97:669–673.

31 McEwen LA, Griffiths J, Schultz K. Developing and successfully implementing a competency-based portfolio assessment system in a postgraduate family medicine residency program. Acad Med. 2015;90:1515–1526.

32 Oandasan I, Holtby A. Trends in resident perspectives of CBME programmatic assessment. Paper presented at Family Medicine Forum 2020 virtual conference, November 4, 2020.

33 Aggarwal M, Holtby A, Oandasan I, eds. Insights on Preparedness for Practice From Family Medicine Longitudinal Survey Data: An Outcomes of Training Project Evidence Summary. Mississauga, ON: College of Family Physicians of Canada; 2022.

34 Ding M, Koppula S, Szafran O, Au L, Babenko O. Mindsets of early-career family physicians trained in competency-based education. PRiMER: Peer-Rev Rep Med Educ Res. 2021;5:39.

35 Schultz K, Griffiths J, Lacasse M. The application of entrustable professional activities to inform competency decisions in a family medicine residency program. Acad Med. 2015;90:888–897.

36 Loeppky C, Babenko O, Ross S. Examining gender bias in the feedback shared with family medicine residents. Educ Primary Care. 2017;28:319–324.

37 Zaki N, Cavett T, Halas G. Field note use in family medicine residency training: Learning needs revealed or avoided? BMC Med Educ. 2021;21:451.

38 Hamza DM, Ross S, Oandasan I. Process and outcome evaluation of a CBME intervention guided by program theory. J Eval Clin Pract. 2020;26:1096–1104.

39 Telio S, Ajjawi R, Regehr G. The "educational alliance" as a framework for reconceptualizing feedback in medical education. Acad Med. 2015;90:609–614.

Appendix 1

Publications Sharing Evaluation Data About CRAFT Implementation in Family Medicine Residency Programs in Canada Used for the Meta-Evaluation of CRAFT

Year	First author	Title	Method	Findings	Assumption supported
2022	Aggarwal[33]	Insights on Preparedness for Practice From Family Medicine Longitudinal Survey Data: An Outcomes of Training Project evidence summary	Longitudinal multicohort survey (Canadian family medicine residency program graduates 2015–2019)	Higher levels of self-reported "Preparedness for practice" were associated with high levels of agreement with multiple elements of Triple C and CRAFT, including "Had an identified person guiding my development" and "Had many informal opportunities for feedback."	1, 2
2021	Ding[34]	Mindsets of Early-Career Family Physicians Trained in Competency-Based Education	Longitudinal survey of one cohort; surveyed at graduation, 1 yr in practice, 3 yrs in practice.	Self-report of goal orientation (labeled "mindset"). Mastery significantly higher than performance at all time points. Authors connect findings to coaching and guided self-assessment in the family medicine residency program (was the CBAS framework).	1
2021	Schultz[30]	Discrepancies Between Preceptor and Resident Performance Assessment: Using an Electronic Formative Assessment Tool to Improve Residents' Self-Assessment Skills	Secondary data analysis of archived FN	Comparison of resident-initiated FN to preceptor-initiated FN: 72% of FN showed agreement on level of supervision needed. Of 28% discordant, 73% showed resident self-assessing lower than preceptors (under-calling performance).	1
2021	Zaki[37]	Field Note Use in Family Medicine Residency Training: Learning Needs Revealed or Avoided?	Secondary data analysis of archived FN	Residents received an average of 32.5 FN over 2 years. Residents are able to form Action Plans to self-direct learning. High variability noted between FNs from rural and urban training streams. A tendency was noted to address "simple" issues rather than complex competencies.	2
2020	Oandasan[32]	Trends in Resident Perspectives of CBME Programmatic Assessment	Longitudinal multicohort survey (Canadian family medicine residency program graduates 2015–2019)	At graduation, > 85% self-reported agree/strongly agree to the following statements: "I was actively aware of my progress during my residency program"; "I could tailor my own learning when learning needs were identified"; "There were many opportunities for feedback on performance"; "I understood what the program expected"; "I had a person(s) guiding my development."	1, 2
2020	Hamza[38]	Process and Outcome Evaluation of a CBME Intervention Guided by Program Theory	Theory-based process and outcome evaluation worked example	Longitudinal evaluation occurred concurrently with implementation by programs. Successful implementation in all 17 programs facilitated by process where findings were used to update program theory of Triple C at CFPC, and were fed back to programs to help with local quality improvement. High-level nature of guidelines facilitated adoption of Triple C by allowing for flexibility in local strategies of implementation. This also contributed to individual programs' sense of autonomy and ownership.	3
2019	Griffiths[29]	Competency-Based Medical Education Implementation: Are We Transforming the Culture of Assessment?	Qualitative grounded theory; interviews with preceptors before full implementation + 3 years post-implementation	Implementation of CBME assessment (CRAFT) resulted in stronger understanding of CBME; perceptions of increased learner engagement in assessment and feedback seeking; perceptions that learners saw feedback as supportive of learning; self-reported higher levels of doing direct observation and providing and documenting feedback (FN); increased perception of value of multiple pieces of assessment data from multiple observers over time for decision making about competence.	1, 2
2018	Ellaway[23]	The Impact of a National Competency-Based Medical Education Initiative in Family Medicine	Realist study (multi-institutional)	Triple C was generally perceived as having had a beneficial effect of teaching, learning, and assessment. Multiple programs reported improvements to direct observation, feedback, documentation, and timelier identification of residents in difficulty (addressing the "failure to fail" problem). Concerns shared about potential overemphasis on struggling residents, which could result in ignoring/overlooking high-performing residents. High-level approach facilitated uptake by allowing for alignment of Triple C to local context—but this also meant a large variation in what elements of Triple C looked like across programs.	1, 2, 3

(Appendix continues)

Appendix 1
(Continued)

Year	First author	Title	Method	Findings	Assumption supported
2018	Ross[27]	Association of a Competency-Based Assessment System With Identification of and Support for Medical Residents in Difficulty	Secondary data analysis of pre-CBAS and post-CBAS archived summative assessments	Significant improvement in identification of residents in difficulty and in addressing difficulties post-CBAS. Significantly fewer residents with poor performance across multiple rotations were seen post-CBAS. Having a continuous advisor helped ensure that any failed rotation evaluations were discussed with the resident.	1, 2
2017	Loeppky[36]	Examining Gender Bias in the Feedback Shared With Family Medicine Residents	Secondary data analysis of archived FN	All preceptors engaged with the assessment system, although female preceptors completed more FN. Overall, preceptors used language in written feedback that emphasized competence rather than character traits.	2
2016	Schultz[26]	Implementing Competency-Based Medical Education in a Postgraduate Family Medicine Residency Training Program: A Stepwise Approach, Facilitating Factors, and Processes or Steps That Would Have Been Helpful	Institutional audit	Increased frequency of direct observation and documented feedback. More efficient and effective identification of "outliers" (residents in difficulty and residents who excel), allowing for earlier and more effective individualized modification of training. Four residents were asked to withdraw from the program after failed remediation; all decisions upheld when appealed due to completeness of assessment information.	2
2015	Schultz[35]	The Application of Entrustable Professional Activities to Inform Competency Decisions in a Family Medicine Residency Program	Institutional audit	10% Increase in numbers of FN after switching from generic FN to EPA FN (6,072 in 10 months 2012–2013 pre-EPA to 6,658 in first 10 months of 2013–2014 post-EPA).	2
2015	McEwen[31]	Developing and Successfully Implementing a Competency-Based Portfolio Assessment System in a Postgraduate Family Medicine Residency Program	Multimethod continuous quality improvement at individual level, program level, and institutional level	Individual level: Residents self-report confidence in ability to identify personal learning needs and are actively aware of their progress during training. Faculty: Improved access to individual resident assessment data facilitates planning individualized learning and early identification of residents in difficulty. Challenges in implementing new assessment balanced by benefits. Program level: Mean FN/resident/year increased from 8 (2010) to 59 (2014). Institutional level: Notable identification and remediation of residents in difficulty, including extensive assessment information and high-quality remediation plans.	1, 2
2012	Ross[25]	Involving Users in the Refinement of the Competency-Based Achievement System (CBAS), an Innovative Approach to Competency-Based Assessment	Grounded theory with participatory action research framework	Resident perception of value of CBAS was influenced by access to feedback that was timely, specific, and frequent. When all 3 elements happened, residents were engaged and motivated. When elements were missing, residents reported CBAS as not worthwhile. CBAS was perceived as useful to learning when there was a variety of assessment data; FN without written feedback were seen as of little value. Residents in sites where preceptors engaged well with the system reported seeing CBAS as beneficial to learning. All residents expressed a need for more faculty development.	2
2011	Ross[24]	Competency-Based Achievement System: Using Formative Feedback to Teach and Assess Family Medicine Residents' Skills	Qualitative participatory action research	PAR used in development of system and for early evaluation data. Educators and program directors: Reported better assessment information, effective at identifying residents in difficulty. Residents: "Cautiously positive" plus need for faculty development and training for residents.	1, 2

Abbreviations: CRAFT, Continuous Reflective Assessment for Training; CBAS, competency-based achievement system; FN, field notes, CBME, competency-based medical education; CFPC, College of Family Physicians of Canada; EPA, entrustable professional activity; PAR, participatory action research.