CANADIAN MEDICAL EDUCATION JOURNAL | REVUE CANADIENNE DE L'ÉDUCATION MÉDICALE

2026, In Press: https://doi.org/10.36834/m4gzta58

ORIGINAL RESEARCH

# Competence committees in context: a multi-site realist evaluation of implementation practices

*Tessa Hanmore,[1] Melissa Andrew,[2,5] Jessica Trier,[7,6] Sussan Askari,[7,5] Maria Hussain,[2,5] Ranil R Sonnadara,[3,4] Anita Acai[3,6]*

*Author information is provided in the back matter of this manuscript

## Abstract

**Introduction**: Competence committees (CCs) are central to competency-based medical education and synthesize multiple data sources to make high-stakes decisions about trainee progression. While widely implemented, CC practices vary considerably across contexts, creating challenges in standardization, workload, and data interpretation. Understanding how contextual factors influence CC functioning is essential for optimizing processes and ensuring defensible decisions.

**Methods:** We conducted a multi-site realist evaluation comparing CC implementation at two academic health sciences centres in Ontario, Canada: Queen's University and McMaster University. Using secondary analysis of existing interview, survey, and observation data, we mapped context–mechanism–outcome configurations across 14 domains, including membership, workload, data use, and remediation. Our analysis focused on identifying shared challenges and site-specific adaptations, and we used realist principles, emphasizing generative causation to guide our approach.

**Results:** Despite differences in implementation strategies and resource structures, findings revealed more similarities than differences. Common contexts triggered mechanisms that produced comparable outcomes across sites. Differences reflected institutional supports: Queen's leveraged academic advisors and educational consultants, while McMaster relied on decentralized resource sharing.

**Conclusions:** Variability in CC practices is not a weakness but an adaptive response to local contexts. System-level priorities include investing in data visualization tools, clarifying CC and Residency Program Committee boundaries, and supporting faculty development to interpret narrative data. Future research should examine developmental orientations and link CC mechanisms to learner outcomes, advancing coherent flexibility anchored in core CC principles.

© HANMORE, ANDREW, TRIER, ASKARI, HUSSAIN, SONNADARA, ACAI

# Les comites de compétences dans leur context : une évaluation réaliste multisite des pratiques de mise en œuvre


## Résumé

**Introduction** : Les comités de compétences (CC) jouent un rôle central dans la formation médicale fondée sur les compétences et synthétisent de multiples sources de données afin de prendre des décisions cruciales concernant la progression des stagiaires. Bien qu'elles soient largement mises en œuvre, les pratiques des CC varient considérablement d'un contexte à l'autre, ce qui pose des défis en matière de standardisation, de charge de travail et d'interprétation des données. Il est essentiel de comprendre comment les facteurs contextuels influencent le fonctionnement des CC afin d'optimiser les processus et de garantir des décisions justifiables.

**Méthodes :** Nous avons mené une évaluation réaliste multisite comparant la mise en œuvre des CC dans deux centres universitaires de sciences de la santé en Ontario, au Canada : l'université Queen's et l'université McMaster. À l'aide d'une analyse secondaire des données existantes issues d'entretiens, de questionnaires et d'observations, nous avons cartographié les configurations « contexte-mécanisme-résultat » dans 14 domaines, notamment la composition des comités, la charge de travail, l'utilisation des données et la remédiation. Notre analyse s'est concentrée sur l'identification des défis communs et des adaptations spécifiques à chaque établissement. Notre démarche était guidée par les principes de l'évaluation réaliste, en mettant l'accent sur la causalité générative.

**Résultats :** Malgré des différences dans les stratégies de mise en œuvre et les ressources disponibles, les résultats ont révélé davantage de similitudes que de différences. Des contextes communs ont déclenché des mécanismes qui ont produit des résultats comparables d'un site à l'autre. Les différences reflétaient les soutiens institutionnels : l'Université Queen's s'est appuyée sur des conseillers pédagogiques et des consultants en éducation, tandis que l'Université McMaster s'est appuyée sur un partage décentralisé des ressources.

**Conclusions :** La variabilité des pratiques des CC n'est pas une faiblesse, mais une réponse adaptative aux contextes locaux. Les priorités au niveau du système comprennent l'investissement dans des outils de visualisation des données, la clarification des responsabilités respectives des CC et des comités de programme de résidence, ainsi que le soutien au développement du corps enseignant pour l'interprétation des données narratives. Les recherches futures devraient examiner l'orientation des comités envers le développement des personnes en formation et établir des liens entre les mécanismes des CC et les résultats des apprenants, afin de promouvoir une flexibilité cohérente ancrée dans les principes fondamentaux du CC.


## Introduction

Competence committees (CCs), also known as clinical competency committees (CCCs) in the United States, are groups of clinicians and educators who review trainee progress and decide on their readiness to advance to the next stage of training.[1] These committees synthesize multiple data sources to make high-stakes decisions about promotion and remediation, aiming to ensure fairness, transparency, and developmental support for trainees.

While CCs are central to competency-based medical education (CBME), their implementation is complex and highly context-dependent. Programs must balance consistency with flexibility, manage workload, interpret large volumes of assessment data, and maintain clarity around roles and processes. Without understanding how local contexts shape CC functioning, it is difficult to identify which practices are transferable and which may benefit from adaptations across programs and institutional contexts.

Existing literature describes CC implementation practices at either single or multiple sites, but few studies systematically compare implementation practices across different sites. Given the multiple challenges that researchers have identified in operationalizing CCs,[2,3,4] identifying key similarities and differences across sites will provide a new lens through which medical educators can collaborate toward effective and efficiently functioning CCs.

Several studies have examined CC implementation practices, focusing on the challenges encountered in operationalizing these committees. Acai and colleagues'[2] three-phase study of CC implementation at an academic health sciences centre revealed challenges related to adapting to program size, optimizing membership, engaging residents, maintaining member capacity, sharing and aggregating data, and developing a clear mandate. Cheung and colleagues[3] described CC implementation practices across different programs at different sites and reported considerable heterogeneity, which they attributed to local adaptations in response to program-specific challenges. A multispecialty ethno--raphic study by Ekpenyong and colleagues[5] reached similar conclusions regarding CC heterogeneity. More recently, Oswald and colleagues[4] discussed challenges in implementing CCs nationally, including standardizing processes and procedures while maintaining flexibility; addressing contextual variability within institutions, programs, and systems; managing finite human and financial resources; providing faculty development and ensuring engagement; and changing the culture of assess--ent. These findings underscore the complexity of CC implementation and the need for frameworks that explain, not just describe, variation. Recent syntheses reinforce the scope of CC heterogeneity and point to secondary functions beyond promotion decisions. Ekpenyong and colleagues'[5] multispecialty ethnography documents wide variation in membership, decision practices, and the informal interpretive work CCCs undertake, and highlights how committees often assume roles in remediation, assessment system improvement, and faculty development. Ekpenyong et al.'s findings are echoed by national implementation analyses that identify system-level challenges in standardizing processes while preserving program flexibility.[4] Together, these studies underscore the need not only to describe variation but to explain how local contexts trigger mechanisms that lead to convergent or divergent outcomes, a question particularly suited to realist evaluation.

A scoping review mapped extensive variability in CC structures and processes and identified persistent barriers—such as insufficient data, interpretation challenges, and bias—alongside facilitators like faculty development and continuous quality improvement.[6] Complementing this, a narrative review synthesizing 84 studies revealed that CCs often assume secondary roles—remediation, faculty development, and assessment system improvement—beyond their primary mandate, yet evidence for long-term outcomes remains limited.[7] Empirical work further quantifies the resource burden of CC operations, estimating annual costs exceeding $400,000 for a mid-sized institution and identifying faculty dissatisfaction with current processes.[8]

A realist evaluation approach offers such a framework for explaining why CCs function differently across context rather than merely documenting variations. By asking "what works, for whom, under which circumstances, and why," realist evaluation moves beyond cataloging differences to uncover context–mechanism–outcome relationships.[9,10,11] This lens is particularly relevant for CCs, where institutional implementation strategies, resource allocation, and program characteristics may trigger different mechanisms yet produce similar or divergent outcomes. Understanding these dynamics can inform more nuanced guidance for CC implementation and support system-level improvements.

Building on the multi-site studies by Cheung and colleagues[3] and Ekpenyong and colleagues,[5] we conducted a cross-site realist comparison of CC implementation at two academic health sciences centres in Ontario, Canada. Our aim was to identify recurring context–mechanism–outcome (CMO) pat--erns, as well as key similarities and differences, to better understand how local contexts shape CC functioning.

# Methods

## Context

We conducted this study using data from two academic institutions in Ontario, Canada: Queen's University in Kingston and McMaster University in Hamilton.

At Queen's University, a small- to medium-sized academic health sciences centre, all 29 residency training programs transitioned simultaneously to CBME in July 2017. There were considerable centralized support and standardization, although programs had some flexibility. All programs received institutional funding to support the initial transition. Additionally, some programs employed educational consultants with Master's or PhD-level training in education.[12] All programs at Queen's Univer-ity implemented a CC in July 2017.

At McMaster University, a medium-sized academic health sciences centre with 58 residency training programs, programs transitioned to CBME on a rolling basis beginning in July 2017. Programs often shared materials with one another to support the transition, but there was less centralized support and standardization than at Queen's University. Some centralized faculty development and resources were provided by the McMaster Postgraduate Medical Education (PGME) Office. Some programs chose to implement CCs prior to their official transition to the new curriculum.

## Realist evaluation framework

Realist evaluations operate on the premise that projects and programs function effectively under specific conditions and that the reactions of various stakeholders share how they function.[9,10,11] We selected a critical realist stance because our aim was explanatory across sites, to identify CMO patterns that transfer beyond a single setting, a complementary aim to ethnographic descriptions. They address questions like what works, for whom, under which circumstances, and why. A key concept in realist evaluation is that programs may produce different outcomes depending on the context. Single-site studies often fail to distinguish between site-specific factors and broader issues. Thus, a realist evaluation framework is well suited to our study.

This study adopts a critical realist epistemological orientation, consistent with realist evaluation principles. Critical realism assumes that social programs function through underlying mechanisms that are triggered in specific contexts to produce outcomes. This stance informed our use of context–mechanism–outcome (CMO) mapping to explain how and why CCs operate as they do, rather than merely describing surface-level variation.[9,10]

We selected a critical realist epistemology and a realist evaluation approach because our aim was explanatory: to identify the CMO configurations that explain why CC practices vary and when particular practices produce comparable outcomes across settings. Ethnographic designs[5] yield rich, in-depth descriptions of processes and interpretive work within single sites, while realist evaluation seeks transferable causal explanations across contexts. Both approaches are complementary; ethnography provides the thick description that can generate hy-otheses about mechanisms, and realist comparison tests and refines those hypotheses across cases. In this multi-site comparison, critical realism allowed us to prioritize generative causation and to map recurring CMO patterns across Queen's University and McMaster University.

Drawing on assumptions from the CC literature and Royal College of Physicians and Surgeons of Canada (RCPSC) guidelines, this realist evaluation aims to examine which contextual factors generalize across contexts and which are unique. We applied realist principles by mapping CMO relationships across 14 domains: CC Structure, CC Membership, CC Member Engagement, Resources, Training, Technology, Information Sharing, Documentation, Data Sources, Decision-Making Processes, Feedback Loop, Remediation and Appeals, Coaching, and Continuous Quality Improvement. Contexts included institutional implementation models and program characteristics; mechanisms captured how CCs organized, reviewed, and interpreted data; outcomes were operationalized as decision processes, meeting frequency, and communication patterns. We iteratively refined CMO propositions through meetings of the multi-institute research team.

## Data collection and analysis

TH and AA developed a data collection template with 14 domains informed by existing literature and our experiences as researchers and educators (Supplementary Material 1). JT, SA, MH, MA, and RS reviewed the template and provided feedback.

TH and AA populated the template using secondary data from their respective institutions. Queen's University contributed interview data from 25 programs and observation data from seven programs' CC meetings. McMaster University contributed survey data from 15 programs, interview data from 16 programs, and observation data from nine programs' CC meetings.[2] We provide a comparative summary of data sources, recruitment strategies, timelines, and analytic procedures for each site in Table 1. This overview illustrates differences in CBME implementation contexts and study designs while presenting key methodological details—such as participant characteristics, data collection processes, and analytic approaches—in a concise format. The table complements the narrative by making cross-site similarities and differences clear and supports the realist evaluation lens applied in this study.

As shown in Table 1, both sites employed multimethod designs involving interviews, surveys, and observations, but differed in implementation timing, recruitment strategies, and resource structures. TH and AA met regularly during the data input process to discuss and refine the template. Once the data were fully entered, the larger team convened to review the template and examine the findings. Each team member identified similarities and differences between the two institutions and captured other, relevant perceptions using short, written notes. The group held several meetings to discuss these perceptions and finalize the findings. In addition to the data-source information summarized in Table 1, Table 2 summarizes key institutional implementation features that provide the contextual foundation for our CMO analysis.

*Table 1. Summary of data sources, recruitment strategies, and analytic procedures by institution*

| Feature | Queen's University | McMaster University |
|---|---|---|
| Implementation Context | All 29 programs transitioned to CBME simultaneously; centralized support; educational consultants engaged. | 58 programs transitioned to CBME using a rolling transition; less centralized support; PGME-level resources. |
| Recruitment | Email + workshops; purposive sampling of CC chairs; proxies (CC members, educational consultants) when chairs unavailable. | Purposive sampling + snowball (Phase 1); later surveys sent to all CC chairs ($n$ = 35). |
| Data Collection Periods | Jan–Jul 2020 (interviews + observations). | Phase 1: May 2017–Feb 2018; Phase 2: Jan 2019; Phase 3: Jun–Jul 2020; observations Aug 2017–May 2020. |
| Data Types | Semi-structured interviews ($n$ = 17; 11 CC chairs, 6 educational consultants); observations of 7 CC meetings. | Surveys (Phase 1: 14-item; Phase 2: 38-item); interviews (Phase 1 & 3); observations of 16 CC meetings across 9 programs. |
| Programs Represented | 25 programs (86% of residency programs); sizes: small (48%), medium (48%), large (4%). | Survey responses from 15 programs; interviews from 16 programs; observations across 9 programs.<br><br>If comparable size categorizations were applied, programs would span small to large; however, size data were not collected systematically. |
| Interview Procedures | Audio-recorded, transcribed verbatim, anonymized; iterative coding; constant comparative approach; NVivo 12 Plus. | Audio-recorded, transcribed, de-identified; Braun & Clarke thematic analysis; double coding; trustworthiness strategies (reflexive journaling, consensus meetings). |
| Observation Procedures | Template-guided field notes; 30 min–3 hrs; 7 observations were conducted. | Detailed notes; longitudinal observation for 2 CCs (3–4 meetings/year); others observed once or twice. |
| Analysis Approach | Iterative coding; codebook refined collaboratively; theoretical sufficiency reached. | Six-step thematic analysis; descriptive statistics for surveys; triangulation with observation notes. |

*Table 2. Key institutional implementation features relevant to competence committees*

| | Institution 1 | Institution 2 |
|---|---|---|
| Number of residency programs | 29 residency programs | 58 residency programs |
| CBME transition model | Simultaneous transition (all programs, July 2017) | Rolling transition (programs phased in from July 2017) |
| Centralized institutional support for CBME | Strong centralized support (funding for transition, PGME-led initiatives) | Limited centralized standardization; PGME provided templates and some faculty development; programs shared resources |
| Educational consultants | Programs had access to funded educational consultants (MEd/PhD) | No centrally funded educational consultants; relied on peer sharing |
| Academic Advisor (AA) program / longitudinal coaching | Centrally mandated AA program (defined role description, standardized meeting cadence and report template, evaluation process; institutional AA handbook available) | AA-like roles present in some programs but implemented variably (no institutional mandate; heterogeneous reporting and cadence) |
| Data management/e-portfolio system | Single institutional system in use (users reported limited aggregation/visualization) | Different institutional system(s) used; similar user-reported limitations in aggregation/visualization |
| Typical CC implementation timing | CCs implemented across programs at CBME launch (July 2017) | Some programs implemented CCs prior to local CBME launch; others later during rolling transition |
| Typical CC size/quorum issues (observed) | CCs typically 5–9 members; some programs larger | CCs typically 5–9 members; quorum challenges noted in larger committees |
| Notable resource differences | Institutional funding for transition, AA handbook, educational consultant access | Decentralized resource sharing; fewer centralized funded roles |

## Researcher positionalities

In her role as an educational consultant, TH worked with five residency training programs at Queen's University both before and during the transition to CBME. She was involved in the implementation and support of the CCs for these programs. AA, in her capacity as a PhD student, collaborated with numerous residency training programs at McMaster University while studying CC implementation and decision-making processes. She worked as an educ-ion scientist at McMaster University at the time of the study.

At the time of the study, JT was the CC chair and CBME lead for the Physical Medicine and Rehabilitation Residency Training Program at Queen's University, while SA was the program director for the same program. MA was the CC chair for Geriatric Psychiatry and a member of the Accreditation Committee at the RCPSC. MH was the program director for Geriatric Psychiatry at Queen's University. RS was an education scientist at McMaster University, has chaired and served on CCs for several specialties across multiple institutions, and chairs McMaster's CBME Evaluation Committee.

We acknowledged that these roles would influence interpretation (e.g., normative preferences for certain structures). To enhance credibility and dependability, we used cross-site triangulation, explicit domain templates, and group sense-making.

TH, AA, and RS are non-clinicians while JT, SA, MA, and MH hold clinical roles.

## Ethics approval

We obtained ethics approval from the Hamilton Integrated Research Ethics Board (HiREB-#4248) and the Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board (File #6015151).

# Results

Our comparative analysis revealed more similarities than differences in CC implementation across the two institutions. Findings are presented under shared challenges and key differences, aligned with the realist evaluation lens applied in this study. Figure 1 provides a visual summary of the CMO patterns identified.

**Context – Mechanism-Outcomes (CMO) Patterns**

| CONTEXT | MECHANISM | OUTCOME |
|---|---|---|
| **Shared Challenges** | | |
| Finite time/capacity; variable size | Distributed review (primary/secondary reviewers | Persistent workload; recruitment challenges |
| Difficulty obtaining sufficient EPA observations | Narrative comments prioritized | Reliance on narrative comments |
| Documented records discordant with firsthand experience | Experiential input used | Contextualized judgements; fairness tension |
| Electronic platforms lacked functionality for aggregating and presenting data | Manual summaries or spreadsheets | Uneven data visibility |
| Unclear CC vs RPC roles for learning plans/remediation | Diverse learning plan practices | Variable timing of learning plans |
| **Institution 1 Specific** | | |
| Mandated AA role | AA summaries | Enhanced continuity in learner coaching |
| Educational consultants | Consultant supported data prep | Improved data preparation and faculty development |
| **Institution 2 Specific** | | |
| Fewer centralized supports | Peer resource sharing | Learning from early adaptors |
| Variably defined AA role | Program level AA interpretation | Inconsistent coaching and data preparation |

AA = Academic Advisor (faculty member providing longitudinal coaching and support to residents)

*Figure 1. Context-Mechanism-Outcome Patterns*

## Shared Challenges

Across both sites, common contexts triggered similar mechanisms that produced comparable outcomes. The findings below are presented according to these recurring CMO patterns including reliance on informal or experiential data and workarounds for data aggregation. We identified these patterns in the following areas:

***CC Membership.*** CCs typically comprised five to nine members, although programs at Queen's programs ranged from 3 to 18 (observations confirmed quorum challenges in larger committees). Member selection processes were usually informal at both sites, based on perceived "fit" and interest in education rather than standardized criteria. Programs reported recruitment difficulties, attributed to workload and limited faculty capacity.

***Workload.*** Workload was a persistent challenge. At McMaster, CC service contributed to mandatory educational points for faculty appointment maintenance. At Queen's, most CC members had no formal protected time. Administrative support (e.g., minute-taking, document dissemination) was present but insufficient to offset the time required for file review and data aggregation, which CC members described as substantial. Workload constraints were frequently linked to recruitment challenges.

***Data use and completeness.*** CCs relied heavily on Entrustable Professional Activities (EPAs) for progression decisions, even when other data sources were available. Programs often struggled to obtain the expected number of EPA observations for each resident, resulting in incomplete datasets for CC review. These gaps made decision making challenging and increased reliance on other forms of evidence besides EPAs, such as other workplace-based assessments completed by the program and, at times, even undocumented data. When EPA data were available, narrative comments were particularly valuable for providing context, as numeric ratings alone offered limited insight.

***Undocumented experiential input.*** CC members at both institutions grappled with the challenge of undocumented data. Although many believed CCs should be "objective" decision-making bodies, committees routinely used undocumented experiential contributions to contextualize documented data, particularly when formal records did not reflect firsthand experience with a resident.

***Technological limitations.*** While the two institutions used different data management systems, both institutions encountered similar challenges. CC members reported difficulty navigating platform features and extracting the information needed for decision making. Many found that the systems lacked functionality for aggregating and presenting data in a way that supported committee discussions. As a result, some programs resorted to alternative methods, such as creating manual summaries or us-

ing spreadsheets, to compile and organize data. These workarounds increased workload and added extra steps to the review process.

***Resident learning plans and remediation.*** At both institutions, the CC's role in developing resident learning plans and facilitating remediation was initially unclear. Over time, policies clarified boundaries between CCs and Residency Program Committees (RPCs), yet variability persisted. Some CCs created individualized plans for all residents; others implemented formal plans only for those not meeting expectations. Some committees hesitated to formalize plans out of concern for resident morale, whereas others used such plans more liberally to signal commitment to support. Several of the differences described below reflect the contrasting institutional implementation contexts summarized in Table 2, particularly with respect to AA programs, educational consultant support, and centralized implementation structures.

## Key differences

Differences primarily reflected institutional contexts rather than program-specific factors:

***Academic advisors.*** AAs, sometimes called Academic Coaches are faculty with a formal longitudinal coaching role with residents. Depending on the program, they may meet regularly with assigned residents, review portfolios and assessment data (EPAs, PPAs, narrative feedback), support resident reflection and wellness, co-create and monitor individualized learning plans, complete structured reports for CCs, and liaise with program leadership when additional supports or remediation are needed. At Queen's University, AAs were implemented as a centrally mandated program, with defined role description, meeting cadence, report templates, and an evaluation process. At McMaster Uni--ersity, AA functions were present in some pro--rams but were implemented variably and without an institutional mandate, resulting in heterogeneity in role expectations, reporting formats, and frequency of contributions to CC deliberations. We observed that a centralized AA program at Queen's University coincided with more consistent longitudinal input to CCs. We did not evaluate the impact of these models on ultimate decision quality or learner outcomes. (See Table 2: Key institutional implementation features—AA and other implementation differences.)

***Educational consultants.*** Both institutions provided centralized support for CBME; however, the Queen's funding model enabled programs to hire Master's/PhD trained educational consultants to assist with documentation and foster communities of practice. McMaster did not formally employ educational consultants.

***Knowledge sharing and institutional support.*** Queen's simultaneous transition to CBME across all programs facilitated shared priorities and greater buy-in to institutional faculty development initiatives. McMaster's rolling transition produced variable buy-in but allowed programs to learn from early adopters. Resource sharing at McMaster occurred through PGME templates and cross-program membership but was less centralized than at Queen's.

Figure 1 summarizes these CMO patterns, illustrating shared challenges across both institutions and site-specific adaptations related to implementation strategies and resource structures.

# Discussion

This multi-site comparison revealed more similarities than differences in CC implementation across two academic health sciences centres. Viewed through a realist lens, common contexts, such as technology limitations, limited time and capacity, and evolving role boundaries, shaped how committees operated. These contexts activated mechanisms including workarounds for data aggregation and reliance on narrative and experiential input. Together, these mechanisms produced similar outcomes across sites, such as heavy dependence on EPA counts supplemented by narrative comments and variability in learning plan practices.

## Why similar challenges persisted

Despite distinct implementation strategies, committees faced comparable workload and data challenges because platform functionality and role ambiguity constrained processes in similar ways. As summarized in Figure 1, these constraints consistently activated the same mechanisms, manual workarounds and experiential sense-making, leading to conver-

gent outcomes across institutions. The reliance on narrative comments across contexts echoes prior research noting that numeric entrustment ratings often lack granularity and may be skewed toward positive assessments.[13] Committees also supplemented formal data with experiential input when documented records were incomplete or discordant, a practice noted in previous studies of CC decision making.[14,15] This interpretation aligns with recent syntheses highlighting CC heterogeneity and the proliferation of local adaptations—such as dashboards and streamlined summaries—to mitigate data and time constraints.[6,7] Costing studies further underscore how platform deficits translate into significant human time and perceived insufficiency of data for defensible decisions.[8]

## Why differences occurred

Where differences emerged, they reflected institutional supports and governance structures rather than program-specific factors. Queen's centralized resources enabled consistent AA roles and access to educational consultants, fostering structured data preparation and coaching. McMaster's rolling implementation and fewer centralized supports produced program-level variation and reliance on peer resource sharing. These observations mirror findings from prior studies that describe how institutional context influences CBME implementation strategies and resource allocation, including centralized planning and consultant support at Queen's and more decentralized approaches at McMaster.[2,3,12,16,17] Despite these contextual contrasts, similar outcomes, such as workload and data reliance, suggest a stable framework underpinning CC functioning across settings. This interpretation reflects broader scholarship advocating for core assessment principles while allowing flexibility for local adaptation.[18,19]

## Implications for practice and policy

The following implications are derived directly from the shared challenges and contextual differences identified in our cross-site comparison. Our findings point to several system-level priorities for improving CC functioning. Clarifying and standardizing role boundaries between CCs and RPCs is essential to reduce ambiguity and variation in remediation and learning plan processes.[1,4] Institutional leaders need to invest in data visualization and workflow capabilities, such as dashboards that integrate EPAs, narrative comments, and milestones, to help minimize manual workarounds and reduce administrative burden.[1,4] Institutions should also budget for CC operations—including protected time, analytics/dashboards, and faculty development—given recent estimates of resource burden[8] and syntheses emphasizing CC secondary roles and continuous quality improvement imperatives.[6,7]

Our data show that a formalized AA program coincided with more consistent longitudinal reporting to CCs at Institution 1. Compared with the more variable implementation observed at Institution 2, this structure was associated with more standardized contributions to committee deliberations. Given we did not evaluate decision quality or learner outcomes, we cannot determine whether one model is superior; however, role clarity and standardized reporting tools may represent important mechanisms worthy of further study.[19,20]

Given the importance of narrative comments and the continued use of undocumented experiential input across both institutions, faculty development in these areas may help support more transparent committee deliberations.[13,14,15]

## Relation to CBD 2.0 developments

The Royal College's Commitment to Action emphasizes enhanced flexibility for CBD program implementation, acknowledging unique obstacles and opportunities across institutions and programs.[20] Our findings substantiate this stance: heterogeneity is not a weakness but a necessary adaptation to local contexts. The goal is coherent flexibility anchored in a common backbone of CC principles—clear mandates, transparent processes, and defensible decisions—with context-responsive mechanisms layered on top.[18,19]

## Strengths and limitations

Strengths of this study include a multi-site, multispecialty perspective and triangulation across surveys, interviews, and observations, with explicit realist CMO mapping.[10,11] Limitations include reliance on secondary data from two institutions and incomplete standardized metrics for decision quality or timeliness; platform capabilities and local policies

also evolve over time, which may affect generalizability.[2,3,4] We also note that both participating sites are medium-sized academic health sciences centres; therefore our findings may not reflect the resource constraints of very small programs or the organisational complexity of very large centres. In addition, both institutions used AAs to varying degrees, and institutions without formal longitudinal coaching structures may experience different CC processes and outcomes. Future work incorporating additional sites and prospective measures of CC outcomes would strengthen the explanatory model and test the transferability of proposed mechanisms.

## Future directions

Future studies could (a) compare the impact of program size and specialty type on CC processes; (b) evaluate academic advisor integration (attendance, report quality, role clarity) as a mechanism influencing decision timeliness and defensibility; and (c) pilot platform enhancements for integrated visualization of narrative and numeric data, assessing effects on workload, consistency, and learner outcomes.

# Conclusion

This multi-site realist evaluation demonstrates that while CC implementation practices share many common challenges across institutions—such as membership, workload, data management, and remediation—the differences that do exist primarily reflect local contexts and resource structures rather than program-specific factors. These findings reinforce that variability is not a weakness but an adaptive response to institutional realities, consistent with the RCPSC's emphasis on flexibility in CBD implementation.

By mapping CMO relationships, our study clarifies which practices are broadly transferable (e.g., consensus-based decision making, reliance on narrative comments to contextualize numeric data) and which may benefit from local tailoring (e.g., academic advisor integration, educational consultant support). Recognizing this balance is critical for optimizing CC functioning and ensuring defensible, learner-centered decisions.

Moving forward, system-level strategies should prioritize technology enhancements for data visualization, clear delineation of CC and RPC roles, and faculty development to support effective use of narra-ive and experiential data. These actions, coupled with coherent flexibility, can strengthen CC processes across diverse settings and advance the goals of CBME nationally.

Author information:

1- Queen's University, Ontario, Canada

2- Department of Psychiatry, Queen's University, Ontario, Canada

3- McMaster University, Ontario, Canada

4- University of Toronto, Ontario, Canada

5- Providence Care Hospital, Ontario, Canada

6- St. Joseph's Healthcare Hamilton, Ontario, Canada

7- Department of Physical Medicine and Rehabilitation, Queen's University, Ontario, Canada

Correspondence to:

Tessa Hanmore

email: tessa.hanmore@queensu.ca

Published ahead of issue:

Nov 24, 2026

*© 2026 HANMORE, ANDREW, TRIER, ASKARI, HUSSAIN, SONNADARA, ACAI; This is an Open Journal Systems article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, reproduction, and adaptation in any medium or format, including for commercial purposes, provided the original work is properly cited.*

Conflict of Interest:

The authors report no conflict of interest.

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector.

Conference Presentation:

This work was presented at ICAM, A Realist, Multi-Site Comparison of Competence Committee Implementation Practices, Quebec City, April 2023.

# References

1. Royal College of Physicians and Surgeons of Canada. *Competence by Design technical guide 3: Competence committees.* 2022. Available from: https://www.bing.com/ck/a?!&&p=cffe62c3ec0908584d7245649392384462eb5a8927b9fa53e62d8b95e2746589JmltdHM9MTc4NDUwNTYwMA&ptn = 3&ver=2&hsh=4&fclid=206776af-ca96-66f0-319f-61f0cbb867e1&u=a1aHR0cHM6Ly9uZXdzLnJveWFsY29sbGVnZS5jYS9jb250ZW50L2RhbS9kb2N1bWVudC9zdGFuZGFyZHMtYW5kLWFjY3JlZGl0YXRpb24vdGVjaG5pY2FsLWd1aWRlLTMtdjEtMS1lLnBkZg [Accessed on Nov 19, 2026].
2. Acai A, Cupido N, Weavers A, et al. Competence committees: The steep climb from concept to implementation. *Med Educ.* 2021;55(9):1067-1077. https://doi:10.1111/medu.14585
3. Cheung WJ, Wagner N, Frank JR, et al. Implementation of competence committees during the transition to CBME in Canada: a national fidelity focused evaluation. *Med Teach.* 2022;44(7):781-789. https://doi:10.1080/0142159X.2022.2041191
4. Oswald A, Dubois D, Snell L, et al. Implementing competence committees on a national scale: design and lessons learned. *Perspect Med Educ.* 2024;13(1):56-67. https://doi:10.5334/pme.961
5. Ekpenyong A, Edgar L, Wilkerson LA, Holmboe ES. A multispecialty ethnographic study of clinical competency committees (CCCs). *Med Teach.* 2022;44(11):1228-1236. https://doi:10.1080/0142159X.2022.2072281
6. Karpinski J, Fantaye AW, Hendry P, et al. Implementation, operationalization, and optimization of clinical competency committees in postgraduate training programs: a scoping review. Acad Med. 2025;100(12):1543-1554. https://doi:10.1097/ACM.0000000000006180

7. Ekpenyong A, Holmboe ES, Govaerts M, Heeneman S. Investigating the roles and impact of clinical competency committees in graduate medical education: a narrative review. *J Grad Med Educ.* 2024;16(6):662-683. https://doi:10.4300/JGME-D-24-00017.1
8. Borges MC, Kennedy MM, Kropf CW, et al. Cost and challenges of clinical competency committees and Milestones assessment. *J Med Educ Curric Dev.* 2025;12:1-10. https://doi:10.1177/23821205251336279
9. Wong G, Greenhalgh T, Westhorp G, Pawson R. Realist methods in medical education research: what are they and what can they contribute? *Med Educ.* 2012;46(1):89-96. https://doi:10.1111/j.1365-2923.2011.04045.x
10. Wong G, Westhorp G, Manzano A, et al. RAMESES II reporting standards for realist evaluations. *BMC Med.* 2016;14(1):1-18. https://doi:10.1186/s12916-016-0643-1
11. Jack K. What is realist evaluation? *Evid Based Nurs.* 2022;25(4):111-113. https://doi:10.1136/EBNURS-2022-103608
12. Pero R, Pero E, Marcotte L, Dagnone JD. Educational consultants: Fostering an innovative implementation of competency-based medical education. *Med Educ.* 2019;53(5):524-525. https://doi:10.1111/medu.13868
13. Ginsburg S, Watling CJ, Schumacher DJ, Gingerich A, Hatala R. Numbers encapsulate, words elaborate: toward the best use of comments for assessment and feedback on entrustment ratings. *Acad Med.* 2021;96(7 Suppl):S81-S86. https://doi:10.1097/ACM.0000000000004089
14. Tam J, Wadhwa A, Martimianakis MA, Fernando O, Regehr G. The role of previously undocumented data in the assessment of medical trainees in clinical competency committees. *Perspect Med Educ.* 2020;9(5):286-293. https://doi:10.1007/s40037-020-00624-x
15. van Enk A, MacDonald G, Hatala R, Gingerich A, Tam J. Not in the file: how competency committees work with undocumented contributions. *Med Educ.* 2024; Advance online publication. https://doi:10.1111/medu.15457
16. Dagnone D, Stockley D, Flynn L, et al. Delivering on the promise of competency based medical education – an institutional approach. *Can Med Educ J.* 2019;10(1):e28-e38. https://doi:10.36834/cmej.43303
17. Stockley D, Egan R, Van Wylick R, et al. A systems approach for institutional CBME adoption at Queen's University. *Med Teach.* 2020;42(8):916-921. https://doi:10.1080/0142159X.2020.1767768
18. Govaerts MJB, van der Vleuten CPM, Holmboe ES. Managing tensions in assessment: moving beyond either–or thinking. *Med Educ.* 2019;53(1):64-75. https://doi:10.1111/medu.13656
19. Pack R, Lingard L, Watling CJ, Chahine S, Cristancho SM. Some assembly required: tracing the interpretative work of Clinical Competency Committees. *Med Educ.* 2019;53(7):723-734. https://doi:10.1111/medu.13884
20. Royal College of Physicians and Surgeons of Canada. *Commitment to action: statement on enhanced flexibility for CBD program implementation.* 2023. Available from: https://web.archive.org/web/20240520181017/https://news.royalcollege.ca/en/newsroom/posts/commitment-to-action-statement-on-enhanced-flexibility-for-cbd-program-implementation.html [Accessed on Nov 19, 2026]