Investigating the Roles and Impact of Clinical Competency Committees in Graduate Medical Education: A Narrative Review

Andem Ekpenyong, MD, MHPE
Eric S. Holmboe, MD
Marjan Govaerts, PhD
Sylvia Heeneman, PhD

ABSTRACT

Background Although Clinical Competency Committees (CCCs) were implemented to facilitate the goals of competency-based medical education, implementation has been variable, and we do not know if and how these committees affected programs and assessment in graduate medical education (GME).

Objective To explore the roles CCCs fulfill in GME and their effect on trainees, faculty, and programs.

Methods We conducted a narrative review of CCC primary research with the following inclusion criteria: all articles must be research in nature, focused on GME and specifically studying CCCs, and published in English language journals from January 2013 to November 2022.

Results The main results are as follows: (1) The primary role of the CCC (decision-making on trainee progress) is mostly described in "snapshots" (ie, focusing on a single aspect of this role at a single point in time); (2) CCCs are taking on secondary roles, some of which were anticipated (eg, remediation, feedback) whereas others were "unanticipated" (eg, use of CCC data to validate trainee self-assessment, predict trainee performance in other settings such as certifying examinations, investigate gender bias in assessment); and (3) Articles briefly mentioned short-term outcomes of CCCs at the level of the trainees, faculty, and programs. However, most studies described interventions to aid CCC work and did not specifically aim at investigating short-term (eg, curriculum changes) or long-term outcomes (eg, improved patient outcomes).

Conclusions CCCs fulfill a range of roles in assessment beyond their intended purpose. A more systematic approach is needed to investigate the outcomes of CCC implementation on GME.

Introduction

The primary role of Clinical Competency Committees (CCCs) in graduate medical education (GME) is to render judgements about trainees' performance against a set of criteria and standards (eg, competency Milestones in the United States) and to make recommendations to program directors about advancement.[1] In addition to this primary role, there has been a recognition of evolving secondary roles that these committees can play in residency training.[2] Although neither a mandate nor rarely made explicit, these secondary roles result from CCCs being privy to "seeing" and synthesizing all their trainees' assessment data and may include remediating trainees, critiquing the quantity and quality of assessment data, providing faculty development, identifying curricular gaps and redundancies, and serving a quality improvement function for the assessment system.[2] To date, however, there has been no synthesis of the existing literature to describe roles actually fulfilled by CCCs and their effect on GME and assessment systems in particular. Without such exploration, we risk not providing CCC members with the time allotment, faculty development, and administrative support needed, thus limiting their ability to ensure that trainees are ready for graduation and capable of providing safe and effective care. Therefore, the purpose of this narrative review is to provide an overview of the published research on the roles CCCs fulfill in GME and their effect on residency training and assessment.

Methods

We chose to conduct a narrative review because they are particularly useful for exploring topics that are complex and under-studied, and they allow the inclusion of a wide variety of studies.[3] Our approach was informed by the recommendations published by Sukhera, which offer guidance on key decisions in the review process.[3] According to Sukhera, narrative reviews are (1) "flexible" in which the initial scope

DOI: http://dx.doi.org/10.4300/JGME-D-24-00017.1

Editor's Note: The online supplementary data contains a visual abstract.

may change through the review process," (2) well suited for topics that are "under-researched," (3) "require a meaningful synthesis of research evidence that may be complex or broad," and (4) "require detailed, nuanced description and interpretation."[3] Our study met each of these criteria. First, although we started the review broadly (eg, including perspectives and policy documents), we realized that we needed to focus on research articles to gather data about how CCCs are actually being implemented and studied. Second, our focus on the roles of CCCs and their associated outcomes is indeed an "under-researched" area in GME. Third, the CCC literature to date is complex (ie, represented by varying implementation approaches).[4] Fourth, we provided a "nuanced"[3] analysis of the descriptions offered in the literature of CCCs as detailed below.

Search Strategy

To our knowledge, the first mandate for CCCs was put forth in January 2013, when the Next Accreditation System went into effect. Thus, we searched PubMed, Scopus, ERIC, and Google Scholar databases from January 1, 2013 to November 1, 2022. During the study, the search was updated twice because there was a delay in our process due to the COVID-19 pandemic and the emergence of pertinent CCC literature. We used "Clinical Competency Committee" as a search term to identify publications where the CCC was mentioned in the title and/or abstract. The search strategy was conducted under the guidance of a librarian at the leading author's institution.

Publication Selection Process

The FIGURE summarizes the search process. One author (A.E.) performed a preliminary review of the search results. One hundred seventeen articles were initially identified and, after removal of duplicates, 111 were selected for "title and abstract review." Two researchers (A.E. and either M.G. or S.H.) independently screened 60 abstracts for inclusion criteria and subsequently resolved discrepancies through discussion. This approach proved to be critical to the richness of our discussions centered on developing definitions for some of the key concepts involved in conceptualizing our work, such as "secondary roles" and "impacts." A.E. subsequently reviewed all remaining abstracts. Following full-text review of selected abstracts, ultimately, 84 articles were included for data extraction.[5-88]

Inclusion/Exclusion Criteria

Inclusion criteria were the following: all articles must be research in nature, focused on GME and specifically studying CCCs (ie, the term CCC was used in the abstract and/or title), and published in English language journals from January 2013 to November 2022. Articles were excluded if they focused on other decision-making groups such as education committees, were written as perspectives, or situated in undergraduate medical education. The authors discussed these criteria at length, opting to focus the review on only GME, as the process for using CCCs in undergraduate medical education is still evolving.

Data Extraction

A data extraction form was created in Microsoft Excel. The categories (codes) for data extraction included demographic program information, description of the primary and secondary role(s) of CCCs, and the effect of these roles at various levels of the assessment system and context (eg, on the trainee, training program, specialty). Using the preliminary coding framework, 3 authors (A.E., M.G., S.H.) independently reviewed 14 randomly selected articles. Preliminary findings were discussed in the entire research team (A.E., M.G., S.H., E.H.) to resolve any discrepancies and to further refine the coding framework. Using the final coding framework, A.E. subsequently reviewed the remaining articles, while the other authors (M.G., S.H., E.H.) "spot checked" the extracted data for accuracy and completeness. Any articles co-authored by one or more members of the research team were reviewed independently by a colleague (Lisa Conforti or Raghdah Al-Bualy) with content expertise in CCCs. TABLE 1 presents a taxonomy of themes capturing both primary and secondary CCC roles.

We were aware of the posited secondary roles mentioned in the CCC guidebook,[2] but we did not know if this would capture the breadth of what was published. Thus, in our taxonomy, we differentiated "anticipated" secondary roles (ie, those that had been mentioned as theoretically possible in the CCC guidebook[2]) versus those that we had not already seen described ("unanticipated"). We used the competency-based medical education (CBME) logic model developed by Van Melle et al to define effects of CCCs on GME training.[89] A logic model is a visual representation of how components of a process interact to effect short- and long-term outcomes.[89] Thus, changes in the assessment system (eg, new rotations, assessment tools) or actions and attitudes of stakeholders resulting from CCCs fulfilling primary and secondary roles were identified as effects or outcomes. We furthermore characterized outcomes as "short-term" (eg, changes that can be measured in 1 or 2 CCC cycles, for example, residents using CCC feedback to create individualized learning plans) versus "long-term" (eg, those changes

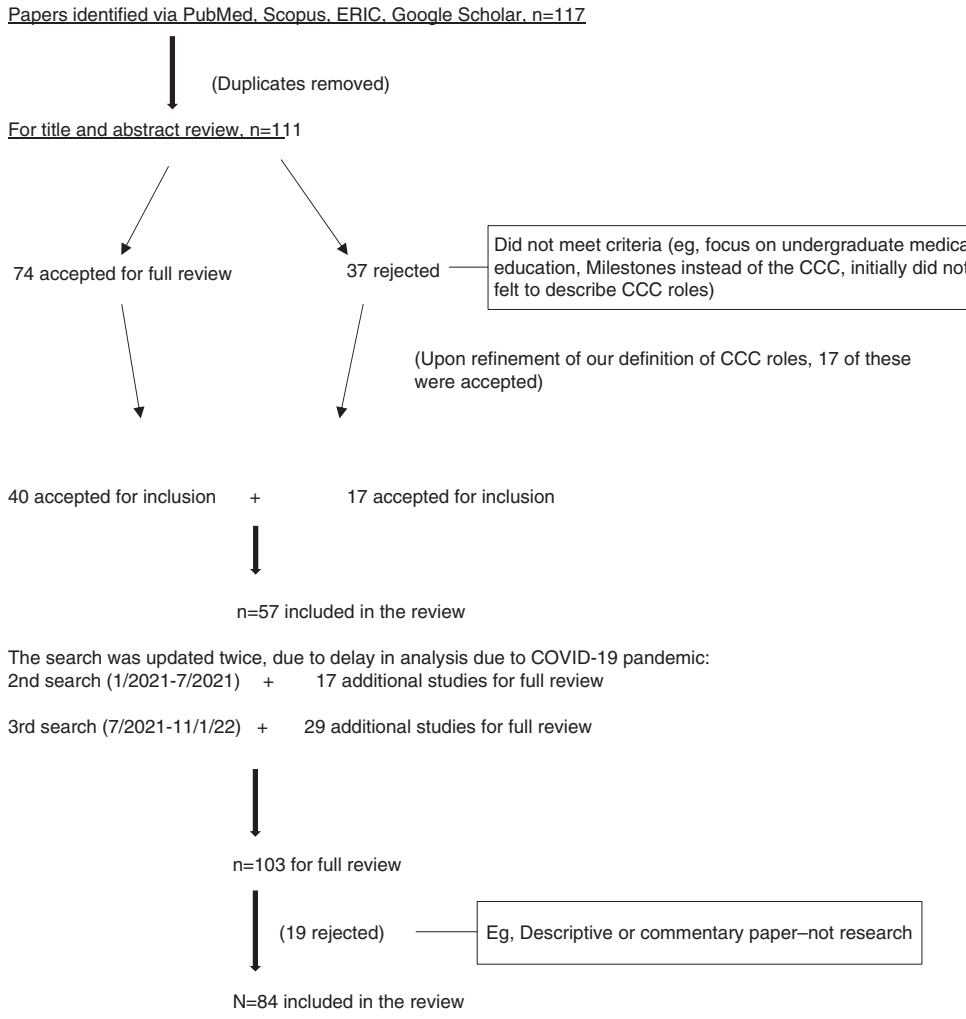

FIGURE
Flowchart of the Literature Search and Study Selection Process
Abbreviation: CCC, Clinical Competency Committee.

that take a much longer time to identify and measure such as linking trainee learning to improved patient outcomes or determining whether CCC ratings influence revisions to the Milestones).

Reflexivity

We had ongoing discussions to determine how our a priori definitions and assumptions of CCC roles and experiential knowledge might influence our interpretation of the data. All authors have published under the broad heading of assessment. A.E. works for the Accreditation Council for Graduate Medical Education (ACGME) and has contributed to the CCC literature. E.H. worked for the ACGME at the time of the study and has contributed to the Milestones and CCC literature. M.G. has been involved in design of assessments in which decision-making relies on CCCs. She also has experiential knowledge of how CCCs work and how they can affect assessment quality. S.H. has been involved in the design and execution of programs of assessment and curricula of undergraduate medical education in which decision-making is done by a committee, and she has experiential knowledge on quality assurance of assessment programs.

Results

TABLE 2 presents an overview of the demographic characteristics of the studies included in our review. Evaluation of the quality of the studies included in our review identified most studies to be original research published in peer-reviewed journals, focusing on specific elements in the process of CCC implementation and decision-making. The majority were

TABLE 1
Summary of the Articles Included in This Review

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
CCC Primary Role			
Scoring and rating patterns			
Beeson et al, 2017[9]	Emergency medicine residents' Milestones ratings	National study reporting descriptive analysis of straight line scoring (SLS)	SLS was found in 20% of emergency medicine programs with no significant difference between the fall and spring scores.
Schwartz et al, 2020[10]	Pediatric residents' assessments	Multi-institutional study to investigate response process validity (Clinical Competency Committee [CCC] process)	Interns from programs that were competency-based and time-variable were found to have the same level of competence as those from traditional programs. CCC decisions may involve undocumented information.
Schumacher et al, 2019[11]	Pediatric residents' CCC ratings	Multi-institution exploratory study to determine association between average Milestones ratings (AMRs) and competencies, and that between AMRs and CCC's supervisory role recommendations	Modest positive correlation between AMRs and supervisory role and moderate to large positive correlation between Milestones ratings for individual competencies and AMRs. Results suggest that performance can be assessed by measuring only 1 or 2 aspects of performance, thus simplifying the process.
Schumacher et al, 2020[12]	Pediatric residents' assessments by CCCs	Multi-institutional prospective cohort study of clinical performance using entrustable professional activities (EPAs)	90% of pediatric residents in the study achieved "unsupervised practice" by graduation for only 8 of 17 EPAs.
Dehon et al, 2015[13]	Emergency medicine resident assessment data (end-of-shift evaluations [ESEs] consisting of emergency medicine Milestones) and other data were used by CCCs	Single institution retrospective analysis of assessment data	Poor correlation between CCC ratings and ESE scores. ESE (which incorporated emergency medicine Milestones) did not distinguish level of training. Faculty rated residents as not achieving a Milestone only 8% of the time. ESEs overestimated performance.
Schumacher et al, 2018[14]	CCC members of pediatric programs	Multi-institutional prospective observational cohort study to investigate the association between the review process CCC members use and their summative decisions	CCC members who provided feedback after CCC meetings were more likely to assign lower Milestones ratings and less likely to recommend supervisory responsibility. Those who reviewed more residents also recommended less supervisory responsibility.
Hamstra et al, 2019[15]	Urology, emergency medicine, and diagnostic radiology (DR) interns' Milestones data	National study to determine whether CCCs are consistent in assigning their ratings	Average ratings decreased over time for 32 of 56 subcompetencies (no change was noted with the others). For instance, in DR 8 of 12 subcompetencies had a downward trend (with effect sizes 0.01-0.03).
Kelleher et al, 2021[16]	Internal medicine residents' narrative data	Single institution qualitative study to compare narrative assessments for struggling and satisfactory performing residents	Thematic analysis of narrative data for the cohorts revealed 2 types of themes: explicit issues centered on resident performance (not paying attention to details, deficits in communication, inability to "see the big picture") and implicit comments based on the assessors description such as the types of phrasing used to differentiate both levels of performance.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Data dilemmas			
Pack et al, 2019[5]	Multispecialty CCCs at single institution (Canada)	Single institution, multispecialty, constructivist grounded theory study to determine how CCCs interpret assessment data	Although CCC members expected to have "high quality" assessment data, they were often in the position of trying to interpret "problematic" data instead. However, their willingness to engage in "effortful" discussions about such data led to meaningful decision-making.
Tam et al, 2020[6]	CCC members, pediatric subspecialty (Canada)	Single institution instrumental case study to investigate the use of previously undocumented data in CCC decision-making	CCC members found formal assessment data lacking and perceived undocumented data (such as personal anecdotes, hearsay, impressions, and contextual information about trainees' situations) to be invaluable to their decision-making.
Schauer et al, 2022[51]	Internal medicine residents' assessment data	Single institution observational study to describe the development of an expected entrustment score for a work-based assessment (ie, observable practice activity)	The expected score is a prediction of performance for a given resident which takes into account the context in which assessment occurs. Analysis of 10 years of data revealed that many residents' observed entrustment scores "oscillated closely" to their expected scores. However, the authors provide examples of residents whose actual scores deviate from the expected score, and in some cases presents an opportunity for CCCs to intervene early.
Roshan et al, 2021[52]	General surgery resident assessment data (narrative comments)	Single institution retrospective analysis to determine the effect of rotation setting on the quality of narrative comments	The McMaster Narrative Comment Rating Scale was used to rate narrative comments generated over approximately a 5-year period. There was variability in the quality of the narrative comments both between and within rotations. Academic sites produced comments that were more corrective and actionable compared to other sites. Urban tertiary rotation comments were of lower quality across the rating categories.
CCC process			
Donato et al, 2016[53]	Internal medicine CCC members	Single institution study of a program's approach to developing a developmental CCC (ie, all residents are discussed)	Authors describe the 2-step presentation process for each resident (using a debate-like set up). Findings included: (1) the importance of faculty development (eg, led to more actionable feedback to residents), and (2) the feasibility of implementing a developmental approach.
Abbott et al, 2021[54]	General surgery residents' end-of-rotation and CCC assessments	Single institution study to determine whether natural language processing (NLP) estimates CCC Milestones ratings	There were 594 end-of-rotation assessments and 97 CCC assessments. The mean for area under the curve was 0.84 for models with non-NLP predictors, 0.83 for those with only NLP predictors and 0.87 for those with both NLP and non-NLP predictors. Authors concluded that NLP could be used to estimate Milestones ratings.
Rozenshtein et al, 2015[55]	Radiology program directors	National observational cross-sectional study via surveys to determine approaches to teaching and assessing radiology residents	Response rates ranged from 35% to 39%. At the time of the study, the development of CCCs was expanding with program directors serving as the chair in most instances. The most used assessment was the end-of-rotation evaluation (96% of respondents). 60% performed preliminary ratings of the Milestones prior to the CCC meeting.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Hauer et al, 2015[7]	Internal medicine residency program directors	Multi-institutional qualitative study conducted via semistructured interviews to determine how CCCs are implemented	Two approaches were revealed in the process of resident assessment: (1) "developmental" (assisting each resident in their learning trajectory) and (2) "problem identification" (focusing on struggling learners). In this study, most of the programs used a problem identification approach.
Lloyd et al, 2020[56]	Program directors, associate program directors in psychiatry, psychiatry residents' assessment data	Multi-institutional study examining the utility of various assessment tools in CCC decision-making	Investigated how CCCs reach decisions (including via the use of actual and mock CCCs) and provided guidelines for these committees. Found end-of-rotation evaluations were the most informative tool. Patient care and medical knowledge were easiest to rate, whereas systems-based practice and practice-based learning and improvement were the most challenging.
Sebesta et al, 2019[57]	Urology program directors	Multi-institutional online survey to determine perceived usefulness of the Milestones	48% did not find the Milestones helpful in assessing their residents. They also felt that the Milestones did not help differentiate clinical performance. 56% of participants had not made any changes to their assessment system based on the Milestones and 80% had not made any curricular updates.
Smit et al, 2019[58]	Pediatric residents and faculty (the Netherlands)	National online survey and conference regarding the implementation of a structured process for rendering summative judgements about resident performance in an EPA-based system	The program was well received by the residents. 75% felt it supported their development with the appropriate level of supervision. 56% received "comprehensive" feedback based on CCC meetings. Faculty also expressed satisfaction with the program (7.3 on a scale of 1 to 10).
Duitsman et al, 2019[59]	CCC members in pediatrics (the Netherlands)	Single institution design-based study to determine how CCCs perform their work via the development of a prototype	Design principles used to develop a CCC could be implemented in practice (eg, structured discussions, a shared mental model, multiple assessment tools). In this study, residents were not satisfied with the feedback they received from the CCC, but this was not addressed prior to the next cycle.
Ekpenyong et al, 2021[60]	CCC chairpersons (multiple specialties)	Multi-institutional cross-sectional web-based survey to identify resources CCCs need	CCC chairpersons supported the notion of institutions offering faculty development to CCC members (81.2%) and workshops for program coordinators (87%). 93.7% did not received protected time for their participation CCC work.
Ekpenyong et al, 2022[4]	CCC members, program directors (multiple specialties)	Multi-institutional ethnographic study of CCC processes	CCC practices were described with respect to 3 themes—membership, CCC roles, and processes such as decision-making and resident presentation to the committee. There was significant variation noted in the approaches used by the CCCs observed.
Ekpenyong et al, 2017[61]	Internal medicine CCC members	Single program mixed methods study to determine how CCCs use different types of assessment data	CCC faculty rated prioritized different types of assessment data: resident rotations, faculty rotation comments, their personal experience with residents (from highest to lowest). They identified 3 types of problems in the assessment process: (1) "design issues" (eg, problematic data or lack thereof), (2) "synthesis issues" (eg, those factors that affect Milestones ratings/decision-making), and (3) "impact issues" (eg, the way the Milestones ratings are used).

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Shumway et al, 2015[62]	Hematology/ oncology fellowship assessment data (EPAs)	Multi-institutional study of the implementation process for EPAs, CCCs	CCCs spent approximately 10 minutes in their decision-making process per fellow. For 5 continuity clinic EPAs, entrustment increased with progression in training. This was the first study of the implementation process for EPAs/CCCs for an internal medicine fellowship. Authors offer recommendations for this process.
Nabors et al, 2017[33]	Internal medicine CCC members, resident assessment data	Single institution study of CCC process outcomes with the implementation of a "rapid deliberation approach"	In the new process, faculty reviewed resident data prior to the CCC meeting and assigned Milestones ratings. These faculty subsequently presented their assigned residents to the CCC and after deliberation consensus was reached. With the new approach, deliberation time was reduced from 25 minutes to 9 minutes per resident and members were more satisfied with the rating process.
Chen et al, 2017[63]	Anesthesiology CCC members	Single institution post-hoc pilot study to evaluate a new process to review resident assessment data	A new process was created using key performance indicators to assist with rating residents on their Milestones. CCC members were involved in the redesign. This CCC assessed "one Milestone at a time, instead of one resident at a time." This process reduced the length of the meeting time (from 8 to 3.5 hours). These CCC members preferred this approach.
Van Osdol et al, 2014[64]	General surgery residents	Single institution study to enable interns to take home call via institution of a curriculum covering multiple core competencies. An innovative Accreditation Council for Graduate Medical Education project	CCCs made the decision as to whether the interns could take home call. Medical knowledge and clinical performance were assessed via nurse mock pages, oral and written examinations. All residents were deemed ready to take home call. Authors outline resident experience including number of calls per shift, duty hours logs, willingness to call for back up from a supervisor, etc.
Kearney et al, 2022[65]	Plastic surgery CCC chairpersons	National survey to determine CCC membership and decision-making process	69.5% response rate. 73% had 1 faculty complete the Milestones ratings and provide feedback to the residents. End-of-rotation assessments and in-service examination scores were used most of the time to determine ratings. 64% felt that Milestones did not facilitate mentorship.
Schumacher et al, 2020[66]	CCC members of a pediatric program	Single institution constructivist grounded theory study to determine how CCC members use resident-sensitive quality measures (RSQMs)	Five dimensions for the use of RSQMs by CCCs were identified: (1) orientation to RQSMs, (2) willingness to use RQSMs, (3) ability to interpret RQSMs, (4) use of RQSMs for assessment decisions, and (5) fairness regarding the use of RQSMs. Authors generated 3 profiles for CCC members based on their engagement with RQSMs: (1) eager, (2) willing, and (3) disinclined.
Goyal et al, 2018[67]	Emergency medicine residents and faculty	Single program study to determine agreement between 3 approaches to decision-making regarding Milestones ratings: (1) traditional or "gold standard" full Milestone assessment (FMA) using all available data, (2) faculty ad-hoc Milestone assessment based on memory (AMA), and (3) resident self-Milestone assessment (SMA)	AMA took less time than FMA but there was agreement between AMA and SMA on only 8 of 23 subcompetencies and between FMA and SMA for 1 subcompetency.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Kou et al, 2022[68]	Pediatric emergency medicine (PEM) fellowship program directors (US and Canada)	Multi-institutional cross-sectional web-based survey to describe if PEM programs entrust their fellows to work independently	There was variability in whether programs allowed their fellows to practice independently. 31 of 68 allowed them to work independently at some point in their fellowship. Programs had certain metrics that had to be met prior to granting this entrustment (eg, Milestones levels attained, number of clinical hours completed).
Schumacher et al, 2019[69]	Pediatric program directors' and CCC members' decisions about residents serving in a supervisory role	Multi-institutional study to explore how program directors "justify" their entrustment decisions about residents' supervisory roles	Authors compared CCC and program director supervisory role ratings. 801 categorizations were made by both. At the 2 time periods studied there was agreement 77% and 91% of the time. In cases of discrepancy, program directors were more likely to assign a lower level of entrustment. Authors list participants' justifications for moving residents either up or down on the entrustment scale. Resident performance played a larger role than their experience when moving residents up on the entrustment scale and the converse when moving them down.
Schumacher et al, 2019[70]	CCC members of pediatric programs	Multi-institutional qualitative study to explore the key factors CCCs use in making decisions about residents being assigned a supervisory role	Via thematic analysis the authors found 4 factors that play a role in CCCs' determination of the type of supervisory role residents are assigned to: (1) trustworthiness demonstrated by the resident, (2) combination of clinical experience and performance, (3) ability to lead a team, and (4) factors beyond the resident such as availability of back-up supervisors for those residents who were deemed "borderline," patient volume, patient acuity, etc.
Schumacher et al, 2020[71]	Pediatric CCC-generated assessment data (ie, expected narrative performance levels [ENPLs])	Multi-institutional study to determine narrative performance levels (per EPA) that denotes readiness for graduation and compare these expectations with actual narratives describing actual performance	In 94.4% of cases, actual initial entrustment scores were the same or higher than the expected performance. For 8 EPAs, ENPLs for graduation and entrustment were the same. For 5 EPAs, the ENPLs for graduation were lower for graduation than for entrustment. Narrative performance at levels 3 and 4 were the primary choices to describe residents at graduation and entrustment.
Stehman et al, 2019[72]	Emergency medicine CCC chairpersons and program directors	National cross-sectional survey to assess performance on the professionalism Milestones	121 programs (65.4%) completed the survey. Faculty shift evaluations (89.7%) and CCC opinions (86.8%) were the most commonly used methods of assessment. Over 95% of graduating residents achieved mastery in accountability (37% of programs) and professional values (42% of programs).
Implementation of the Milestones/EPAs			
Schumacher et al, 2021[73]	Pediatric program leaders and CCC members	Multi-institutional constructivist grounded theory study using interviews to describe the experience of implementing EPAs	Themes generated regarding EPA implementation: (1) recognizing the EPAs as an acceptable framework in "alignment" with current approaches, (2) identifying curricular gaps, (3) recognizing the importance of faculty participation in the process, and (4) recognizing how CCC structure and process impacts the process. Within each of these 4 themes there were factors that facilitated or inhibited the EPA implementation process.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Yaghmour et al, 2022[74]	CCC chairs, residents, program directors (emergency medicine, internal medicine, family medicine, pediatrics)	Multi-institutional qualitative study using interviews and template analysis to explore experience in implementation of Milestones	Authors found 3 stages of Milestones implementation: (1) "resource intense" (eg, lack of engagement of residents and faculty, negative perception of the Milestones), (2) "transition" state (eg, mentions of some improvements in the process), and (3) "fine tuning" stage (eg, faculty development, continuous program improvement).
Balach et al, 2022[75]	Orthopedic surgery program leaders (eg, program director or CCC chairperson) and residents	National qualitative study of orthopedic surgery programs to assess implementation of their Milestones	57 (23%) programs participated in phone interviews. 2 themes emerged: implementation and impact of the Milestones via which benefits and challenges were identified.
Nabors et al, 2013[76]	Internal medicine faculty, program leaders, residents, CCC members	Single institution study describing the implementation of a Milestones-based evaluation process.	Faculty and residents agreed (somewhat or strongly) that the Milestones promoted a shared mental model of expectations at various points in training and enhanced feedback efforts. Faculty and residents also thought that Milestones promoted "fairness and uniformity." 8 of 12 faculty indicated they would recommend a Milestones-based evaluation system to other institutions.
Maranich et al, 2022[77]	CCC members from transitional year programs from 7 specialties (neurology, emergency medicine, ophthalmology, internal medicine subspecialty, pediatrics, internal medicine, and pathology)	Multi-institutional study using a constructivist lens performed a thematic analysis of interview data to explore participants' understanding and application of the Milestones	Found 5 themes: (1) lack of formal training regarding their role as CCC members, (2) perspective of Milestones as a helpful tool for assessing residents, but not referencing the other uses of Milestones (eg, for program evaluation or quality improvement), (3) the existence of a process to convert assessment data into a Milestone rating, (4) use of both formal and informal narrative comments in reaching judgements, and (5) assigning residents an "average" rating in the absence of data (eg, for the systems-based practice Milestones).
Hanson et al, 2023[78]	Advanced gastroenterology/ hepatology fellows and faculty	Multi-institution study to describe the EPA implementation process	Entrustment increased over the course of a year. Overall, viewed as positive by fellows and faculty.
Lewis et al, 2021[79]	Program directors, associate program directors, CCC chairpersons (19 specialties and subspecialties in a children's hospital)	Single institution collective case study to describe the Milestones implementation process	Thematic analysis of meeting notes revealed varied assessment methods and tools to address the 6 core competencies. Three categories of assessment methods emerged: (1) direct (eg, direct observations, in-training examinations), (2) indirect (eg, rotation evaluations), and (3) multisource (eg, 360-degree assessments). Among other issues, the primary challenge identified by participants was the need for alignment between the assessment methods/tools and the Milestones.
Yamazaki et al, 2022[80]	Obstetrics and gynecology (OB/ GYN) residents' Milestones ratings	National study to investigate the internal structure validity of the OB/GYN Milestones	Milestones ratings increased as residency training progressed. Apart from the 6 general core competency domains, CCCs further identified 3 factors of the patient care competency.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Identification of struggling residents			
Schumacher et al, 2018[81]	Pediatrics program directors and CCC members	Multi-institutional, qualitative study using survey methodology and thematic analysis to examine how struggling residents are identified	Found 2 groups of themes by which a resident can be identified for concern: (1) pathways by which a resident reaches a threshold (eg, accumulation of isolated data points, narrative comments from rotation evaluations, developmental), and (2) how CCCs and program directors make decisions about performance concerns (eg, CCC used norm and/or criterion referenced approaches and made judgements as to the quality of the assessment data reviewed including the type of rater [eg, "hawk or dove"])
Conforti et al, 2018[82]	Program directors, residents, and CCC chairs of neurosurgery programs	Multi-institutional grounded theory study to examine the process of using the Milestones to assess neurosurgery residents	Exploration of the Milestones implementation process led to 2 types of responses: (1) "outcomes" (eg, identifying struggling trainees earlier, curriculum revision, continuous quality improvement), and (2) "facilitators" (eg, clarified the core competencies) and "barriers" (eg, not being accepting of the Milestones as a framework).
Dashboards/displays/apps			
Yen and Lewis, 2022[24]	Radiology residents, faculty, and program coordinator	Single institution survey analysis of the implementation of a web-based resident management system	The online platform was well received by residents (82% agreed or strongly agreed that it was a useful addition, 53% agreed that it was easy to navigate). Faculty had a similar positive response. Program coordinator preparation time (for CCC meetings) was reduced by ~90% and that of faculty by at least 50%.
Stahl et al, 2020[25]	Residents and faculty, general surgery, emergency medicine, hospital medicine	Single institution, multi-specialty study of EPA implementation via mobile app. Feedback was obtained from participants about their experience using the app.	Resident self-assessments and faculty assessments were compared (66% were matched, ie, resident and faculty data on the same assessment). Assessments increased after resident and faculty development sessions. Residents and faculty found the app easy to use but identified barriers and offered suggestions for improvement.
Friedman et al, 2016[26]	CCC members, internal medicine	Single institution study of the development and implementation of a dashboard to compile both quantitative and qualitative assessment data	Dashboards were created in Excel for each resident (5 minutes per resident). CCC members took 30 minutes per resident assessing their data. Dashboards facilitated a structured and standardized process to CCC decision-making.
Almufarrej et al, 2022[27]	Plastic surgery faculty, residents' data, CCC members	Single institution prospective pilot study to determine feasibility of the use of a smartphone-based evaluation system	Pre- and post-surveys were conducted. CCC members were also interviewed to determine the effectiveness of feedback obtained via the new system. CCC scored the new app highly (4.65 of 5, compared to a rating of 2.82 of 5 with the previous assessment tool).
Dougherty et al, 2022[28]	CCC members, general surgery	Single institution implementation study of a "resident scorecard" to facilitate CCC decision-making	Scorecard is a single page document containing both quantitative and qualitative assessment data. Due to the implementation of the score cards, total time of the CCC meeting was reduced from 126 to 106 minutes. Time spent per page increased from 1 to 5 minutes (although total pages per resident decreased from 9 to 1).

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Johna and Woodward, 2015[29]	General surgery resident assessment data	Single institution study using a newly developed dashboard to better align existing curricula goals and objectives with the Milestones to enhance CCC deliberations	Developed a dashboard using Microsoft Excel linking data from assessments to each of the Milestones. This reduced the time spent per resident from 45-60 minutes to 5-10 minutes. The dashboard also included the use of radar plots.
Input from allied health professionals			
Regan et al, 2018[30]	Emergency medicine residents' CCC data, end-of-shift (EOS) and end-of-rotation (EOR) data	Single institution study to determine if 2 new assessment methods, EOS and EOR, facilitate CCC decision-making.	EOS and EOR were collected from doctors (d) and nurses (n). EORn and EORd scores were correlated with CCC scores. The strongest correlation was between CCC scores and EORd (r 0.71 to 0.84). EOSn and EORn were not as strongly correlated with CCC scores (r 0.49 to 0.62) and (r 0.4 to 0.73) respectively. Authors also found that levels of proficiency assigned by the CCC were highly correlated with the levels that were assigned in the previous 6 months (suggesting a possible bias by these previous decisions).
Bedy et al, 2019[31]	Emergency medicine resident data, CCC members	Single institution comparison of pharmacists' and CCC's ratings.	Pharmacists assessed residents on subcompetency patient care (PC)-5 ie, "pharmacotherapy." Authors found 100% agreement between the pharmacists' ratings and those of the CCC.
New assessment tools/examinations/curricula			
Mamtani et al, 2015[18]	Emergency medicine residents and CCC members	Single institution study using survey methodology to elicit perceptions from residents and CCC members regarding introduction of a novel curriculum ("EM debates") for systems-based practice and practice-based performance improvement content	71% of participating residents the curriculum improved their performance. 100% of those that led a debate felt they could use an evidence-based approach to a particular topic. CCC members felt that as a result of the new curriculum they could better assess performance on the associated subcompetencies.
Cooney et al, 2016[19]	Plastic surgery resident assessment data	Single institution study of the implementation process of a web-based tool (comprehensive observations of resident evolution "CORE")	CORE distinguished novice from experienced residents. 38% of residents reported more timely feedback about their operative performance.
Mount et al, 2014[8]	Internal medicine residents and faculty	Single institution study of the development of an oral examination for use by the CCC	Validity evidence was obtained. Faculty spent less than 0.1 annual full-time equivalent on this assessment. 54 (86%) residents passed it on their first attempt. The posttest survey of residents and faculty revealed that the examination was perceived as fair.
Bryant, 2021[22]	Pathology residents, faculty, pathologists' assistants, CCC members at a single program	Single institution study using survey methodology to determine feasibility of an EPA for frozen section training	100% of faculty and CCC members agreed that the EPA was easy to understand. 80% of CCC members thought it would help them determine residents' readiness for taking call. 75% of residents found it easy to use and helpful for understanding the frozen section process; 50% thought it provided useful feedback.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Rassbach and Blankenburg, 2018[23]	Pediatric residents and faculty	Single institution study to determine if a coaching program could improve the quality of feedback residents receive and faculty members' confidence in providing feedback	99% of residents agreed or strongly agreed that their coaches were well trained in providing feedback. Coaches reported more confidence in providing feedback and teaching other clinical skills compared to non-coaches. Non-coaches also may have benefited from this program as they reported providing equal or more feedback.
Secondary CCC roles			
Anticipated			
Feedback to trainees and programs			
Pack et al, 2020[32]	CCC members, multispecialty (Canada)	Multispecialty, single institution constructivist grounded theory study to identify CCC roles	CCC members were observed and interviewed. CCCs went beyond their primary role to identify and offer solutions to problems at both the level of the residents and the program. A portion of the work involved occurred outside of formal meeting time.
Remediation			
Warburton et al, 2017[34]	Internal medicine residents' assessment data	Single institution study of a remediation process (ie, Early Intervention Remediation Committee)	CCC referred learners to the remediation committee. All identified issues (eg, medical knowledge, professionalism, clinical reasoning, communication skills, and organization) were resolved. Required investment of ~45 hours per learner. The process was well received by learners and faculty.
Blau et al, 2021[35]	Plastic surgery residents	Single institution case study on implementing their CCC and its various roles	CCC provided residents with feedback and created individualized learning plans—including making reading recommendations and monitoring residents' progress.
Ketteler et al, 2014[83]	General surgery CCC members	Single institution study of Milestones and CCC implementation via the inclusion of "competency champions" with expertise in each of the core competencies	After one year, residents and faculty were more aware of the importance of the core competencies and the remediation. There was also an improved ability by faculty to evaluate the residents and provide feedback. The competency champions provided a monthly report to the CCC on the progress of residents requiring remediation. These competency champions also made recommendations as to whether such residents completed the remediation process.
Address curricular gaps			
Klutts et al, 2015[84]	Pathology resident assessment data	Single institution study describing the process of developing their CCC	Authors describe in detail their CCC process including unique features such as faculty assessing all residents on only those Milestones which the faculty member has expertise in. They also discuss the impact of the CCC on their program eg, residents rated themselves lower on the systems-based practice Milestone that includes laboratory management, thus information on laboratory management was added to the curriculum and pathology residents were asked to attend hospital safety meetings (to address safety and quality issues with respect to laboratory management).

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Unanticipated			
Validation/calibration of self-assessments			
Lyle et al, 2016[36]	General surgery residents	Single institution study to determine if there is a correlation between residents' self-assessments and CCC Milestones ratings	Major differences were defined as >0.5 on a 4-point scale. 7% had a mean overall difference of >0.5, 6 had a difference of ≤0.5. Self-assessment scores were lower than those of the CCC (median difference of -0.06 (-0.25 to 0.16, *P*=.041). Negative self-assessments were more common for medical knowledge. Subgroup analysis by sex, postgraduate level, and time of academic year provide notable areas for further investigation.
Watson et al, 2017[37]	General surgery residents	Multi-institution prospective comparison of CCC ratings to residents' self-assessments	Major differences were defined as >0.5 on a 4-point scale. 13% self-assessed >0.5 higher than the CCC and 13% self-assessed lower than the CCC. Difference for various subgroups (eg, level of training, sex, type of program) were reported. Overall found good concordance between resident self-assessment and CCC scores.
Tichter et al, 2016[38]	Emergency medicine residents, CCC members	Single institution retrospective study to determine agreement between resident self-assessments and CCC ratings	There was moderate agreement between residents' self-assessments and their CCC ratings. No change between the mid and end-of-year reporting cycles.
Ross et al, 2016[39]	Anesthesiology residents' self-assessments and CCC ratings	Single institution study to determine if there is a correlation between residents' self-assessments and CCC ratings	Found high correlation between resident self-assessment and CCC scores in anesthesiology (with correlation coefficients for each Milestone score versus postgraduate year [PGY] level from 0.8-0.95).
Srikumaran et al, 2019[40]	Ophthalmology residents' self-assessments and CCC ratings	Single institution comparison of CCC and resident self-assessment scores	No significant correlation between resident's self-assessment and CCC ratings. For residents' first assessment, CCC and self-assessment score were only strongly correlated for 4 of 24 Milestones—with less difference for medical knowledge and systems-based practice. Differences between self-assessment and CCC scores did not lower significantly with further assessments.
Meier et al, 2016[41]	General surgery residents' self-assessments and CCC ratings	Single institution comparison of CCC and resident self-assessment scores	Compared residents' self-assessment with those of both their faculty advisors and CCC. High correlation noted between the 3 groups of raters. Makes the argument for faculty and residents having more of a stake or ownership of the assessment process.
Chow et al, 2019[42]	Plastic surgery residents' self-assessment and CCC ratings	Single institution comparison of CCC and resident self-assessment scores	Found concordance between residents' self-assessments and CCC evaluations. The exception was second-year residents who self-assessed higher on medical knowledge and patient care. The authors concluded that residents are capable of accurate self-assessment. They also used areas of discordance to make curricular improvements.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Exploration of gender bias in assessment			
Kwasny et al, 2021[43]	General surgery residents' Milestones ratings	Single institution retrospective study to determine if there are gender differences in Milestones ratings (for both CCC ratings and self-assessments)	Examined found no significant difference in CCC scores of male and female residents and scores improved as residents advanced in training.
Santen et al, 2020[44]	Emergency medicine residents' Milestones ratings	National study to compare male and female Milestones ratings in emergency medicine	Did not find major differences in EM male versus female residents' Milestones scores across most competencies (small differences noted in patient care Milestones). Raises the issue of CCC possibly neutralizing biases in the assessments they receive.
Zuckerbraun et al, 2020[45]	Pediatric emergency medicine fellows' Milestones ratings	National retrospective cohort study to determine if there are gender differences in Milestones ratings	There was no difference for male and female fellows in average Milestones scores across the 6 core competencies. By subcompetency females scored higher on PC3 (providing transitions in care) and ICS2 (understanding emotion and human response). This was significant, but not meaningful.
Hauer et al, 2021[46]	Internal medicine residents' CCC ratings	National retrospective study to determine if there are gender differences in end-of-year Milestones ratings for patient care (PC) and medical knowledge (MK) competencies	For the MK Milestones there was no significant difference between scores for male and female residents at PGY-1. MK ratings were higher for PGY-2 and PGY-3 men compared to women. For PC, men and women had similar scores at all PGY levels.
Athy et al, 2021[47]	Pathology residents' assessments	Single program retrospective study comparing CCC and resident self-assessment	No difference between male and female self-assessment scoring. PGY-1s overestimate their skills and PGY-4's ratings were lower than the CCCs'. Self-assessment scores were lowest on systems-based practice and highest on professionalism Milestones. Increase in in-service examination scores correlated with increases in self- and CCC-scoring.
Predicting performance in other settings			
Turner et al, 2016[48]	Anesthesiology residents' assessments	Multi-institutional retrospective study to determine the effect of performance deficiencies on graduation and board certification rates	CCC scores and recommendations were used to predict residents who might have problems on their certifying examination and graduation. The graduation rate among those who had a performance deficiency was 93% (vs 99% in those without a deficiency). However, this was 55% if there was a deficiency in an "essential attribute" (eg, being honest, respectful, ethical, etc) 89% of those with a deficiency attained board certification compared to 99% of those without a deficiency.
Kimbrough et al, 2018[49]	General surgery residents' Milestones ratings and American Board of Surgery In-Training (ABSITE) scores	Single institution study to determine if mid-year CCC ratings could predict ABSITE scores	"Minimal to small" positive correlations were observed between Milestones and ABSITE percentile scores (r=0.09-0.25) with a stronger correlation between Milestones scores and percent correct scores (r=0.65-0.76). The mid-year MK1 Milestone scores contributed significantly to this effect.

TABLE 1
Summary of the Articles Included in This Review (continued)

Studies (Authors, Year)/ Themes	Participants/Data Sources	Type of Study/Methods	Key Finding/Contribution to CCC Role
Rebel et al, 2019[21]	Anesthesiology residents	Single institution comparison of CCC Milestones ratings and those of an objective structured clinical examination (OSCE)	CCC Milestones ratings were higher than those from the OSCEs—this effect was highest with junior residents. However, ratings improved over time for both CCC and OSCE ratings. There was also an expectation bias in the CCC ratings (primarily involving senior residents).
Cullen et al, 2022[20]	Multispecialty (16 residency, 1 fellowship)	Multi-institution to determine if situation judgement test (SJT) scores can identify professionalism issues and if they are correlated with CCC professionalism ratings	SJT scores correlated positively with CCC professionalism Milestones ratings (r=0.21 and 0.14 for mid-year and year-end, respectively).
Validity evidence for Milestones/EPAs			
Bienstock et al, 2021[85]	OB/GYN residents (1184 residents from 240 programs)	National study to evaluate the correlation between CCC ratings and board scores	The strongest association was seen for all 7 subcompetencies of medical knowledge. Correlations were stronger with each subsequent Milestones rating period.
Mink et al, 2020[86]	Pediatric fellows' assessment data	Multi-institution study to determine agreement between fellowship program directors and CCC entrustment decisions	For each of 6 EPAs, there was strong correlation between fellowship program director and CCC entrustments ($P<.001$)—although not as strong when the fellowship program director was not a CCC member.
Mink et al, 2018[87]	Over 200 pediatric fellowship programs	Multi-institutional study using a modified Delphi approach was used to develop scales for common EPAs and demonstrate validity evidence	Supervision scales were used by CCCs and fellowship program directors to assess fellows. Progressive entrustment was demonstrated for each EPA.
Evaluation of curricular elements			
Frey-Vogel et al, 2016[50]	Pediatric interns	Single institution study to determine the correlation between summative simulation cases and CCC scores	There was a statistically significant relationship between summative and CCC scores. An increase of 1 point in summative scores was associated with a 0.2 increase in CCC ratings. Performance on formative and summative cases was not correlated (no *P* values were significant).
Mikhaeil-Demo et al, 2021[88]	Graduating neurology residents	Multi-institutional cohort study to assess status epilepticus management and correlation Milestones ratings	For simulation-based assessments regarding status epilepticus management (SE), simulation-based scores were compared with resident's clinical experience (including CCC scores). Graduating residents scored poorly on simulation-based assessments (average SE checklist score was 60% correct); however, their end-of-rotation scores were higher (at level 4.3 of 5 for epilepsy Milestone and 4.4 of 5 for management and treatment). Concluded that end-of-rotation scores should not be used alone in judging Milestones performance for such skills.

Note: All of the articles in the table were included in the study. There may be articles that overlap multiple categories; however, in order to list each article once, the authors associated each article with a particular theme.

single-institution, single-specialty studies conducted in the United States involving CCC review of trainees at the residency level. See TABLE 1 for a brief description of each of the studies included in the review.

CCC Primary Role and Potential Associated Effect

The primary role of the CCC is to assess trainees and make recommendations to program directors regarding developmental progress.[1] We found that most articles did not describe the primary role in its entirety (ie, from review of the assessment data to making a judgement and communicating it to programs and trainees). Rather, the focus was on various aspects such as trying to make sense of difficult data,[5] the use of undocumented assessment data,[6] role conceptualization,[7] or implementation of new assessment tools.[8] See data for themes and examples. Some articles focused on the process of Milestones ratings and patterns in performance scoring.[9-16] An ethnography of CCCs furthermore revealed marked variability in the approaches CCCs use to conduct their work, such as the strategies used in decision-making and types of assessment data used.[4] We also found little inquiry as to how decisions about trainee performance are actually made. The CCC process (before, during, and after meetings) as described by Al-Bualy and colleagues outlines multiple steps during CCC work.[17] Our review reveals that still very little is known about how these steps are actually performed.

TABLE 2
Demographic Data of the CCC Literature

Characteristics	n (%) N=84
Country	
USA	77 (91.6)
Canada	4 (4.7)
Netherlands	2 (2.3)
Multinational	1 (1.1)
Institution	
Single	46 (54.7)
Multi-institution	25 (29.7)
National	13 (15.4)
GME level	
Residency	78 (92.8)
Fellowship	6 (7.1)
Specialty	
Pediatrics	17 (20.2)
General surgery	11 (13)
Internal medicine	10 (11.9)
Pediatrics/emergency medicine	10 (11.9)
Emergency medicine	9 (10.7)
Plastic surgery	5 (5.9)
Anesthesiology	4 (4.7)
Pathology	4 (4.7)
Obstetrics/gynecology	2 (2.3)
Radiology	2 (2.3)
Urology	1 (1.1)
Neurosurgery	1 (1.1)
Psychiatry	1 (1.1)
Hematology/oncology	1 (1.1)
Ophthalmology	1 (1.1)
Neurology	1 (1.1)
Orthopedics	1 (1.1)

Abbreviations: CCC, Clinical Competency Committee; GME, graduate medical education.

Although we did not find any studies that explicitly aimed to investigate the effect of CCC implementation on programs and assessment systems, there were multiple examples of "CCC-inspired changes" aimed to facilitate CCC work such as instituting new (1) curricula,[18] (2) assessments,[19-23] (3) ways of organizing the assessment data,[24-29] and (4) assessors.[30,31] Beyond these CCC-inspired changes, there were also some mentions of "short-term" outcomes involving the trainees, program, or CCC itself. For instance, at the level of the trainees, one program implemented a web-based direct observation assessment tool with 38% of residents subsequently reporting more immediate feedback on their operative skills.[19] At the level of the program, Mamtani et al describe development and implementation of a curriculum and instructional strategies to support assessment of Patient Safety and Quality Improvement Milestones.[18] Similarly, Pack and colleagues describe CCCs providing training programs with input for development of new learning experiences, based on what CCCs count as evidence for their decision-making and to fill gaps within the assessment system.[32]

Lastly, there were a number of instances of short-term outcomes on the CCC itself. In some cases, authors were able to report a resultant change to their process. For instance, the score cards used in one general surgery program to consolidate all relevant assessment data to a single page card reduced the length of the CCC meeting from 126 minutes to 106 minutes, yet the time available to actually discuss each resident increased from 1 minute prior to using the score cards to 5 minutes after implementation.[28] Nabors and colleagues described reduced deliberation time during CCC meetings following a change in the assessment process (ie, CCC members assigning their ratings prior to the CCC meeting).[33] Other studies reported the effect of including allied

health professionals as assessors on the robustness of CCC decision-making. For example, Bedy and colleagues surveyed emergency medicine CCC members and found that they perceived pharmacists' assessments as useful to their judgement of resident performance.[31] In another study, nurses' end-of-rotation assessments of emergency medicine residents did not correlate with final CCC proficiency levels; however, their descriptive comments were thought to be "invaluable" for identifying areas for improvement.[30] Pack and colleagues found that through the CCCs struggle with problematic evidence, members unpacked their assumptions about the source of the data used and reflected on how they perceive certain types of data as useful evidence.[5] All of these examples suggest the potential effect of CCCs at multiple levels in residency training and assessment, not only on trainees, curricula, and assessment programs but also on the CCC members themselves, for example through development of shared mental models around decision-making processes, including usefulness of assessment data.

CCC Secondary Roles and Potential Associated Effect

We found only one study that was specifically designed to explore CCC secondary roles.[32] In this study by Pack and colleagues, CCCs deliberately contributed to ongoing evaluation of the curriculum and assessment in residency training and provided meaningful feedback to program directors about program limitations, thereby documenting an impact on CCC members, other faculty, and the program.[32] As the vast majority of articles included in our review were not designed to explore CCC secondary roles, we extrapolated the secondary role of the CCC based on the narrative provided by the authors. For instance, some CCCs went beyond identifying struggling residents to assume other responsibilities in the remediation process. Warburton and colleagues describe their CCC not only identifying struggling learners but also referring them to the Early Intervention Remediation Committee.[34] In another setting, CCC members helped create individualized learning plans and monitored residents' progress.[35] In these descriptions of secondary roles, the associated outcomes were not reported.

In addition to these "anticipated" secondary roles, we also found a myriad of secondary roles which had not been ascribed to CCCs previously and denoted these as "unanticipated" secondary roles (see TABLE 1 for themes and examples). The very presence of a CCC, and the data generated by these committees, allowed for "CCC data facilitated research" (ie, use of CCC data to investigate a variety of assessment-related topics). Several studies, for example, investigated the correlation between resident self-assessment and CCC scores.[36-42] In various specialties and settings, CCC data was used to check for gender-based differences in Milestones ratings[43-47] (ie, for evidence of gender bias in assessment). In those studies that did not find any significant gender-based differences,[43-46] some authors hypothesized that CCCs may serve to neutralize biases in the assessments they receive.[46] CCC data were also used to investigate correlations between Milestone ratings and performance in other settings, such as certifying or in-training examinations.[48,49] Lastly, CCC data was used to support the evaluation of curricular activities or changes. For instance, Frey-Vogel and colleagues compared pediatric interns' performance scores in a simulated setting with CCC performance ratings.[50] In all of the examples mentioned above, CCC data was seen as an instrumental resource for investigating a variety of assessment-related questions that arise in GME.

Discussion

To our knowledge this is the first narrative review of the CCC literature focusing on CCC roles and the effect of these committees on residency training and assessment. Our work revealed 3 key findings: (1) most CCC studies address components of the primary role but do not explore the entire process; (2) CCCs fulfil secondary roles that we did not anticipate; and (3) despite the myriad of short-term "CCC-inspired changes," there is a lack of documented long-term outcomes.

By revealing the lack of a comprehensive approach to studying their primary and secondary roles, our review illustrates the complex nature of the work of CCCs. Not only are we still trying to understand how CCCs accomplish their intended role, but our findings also suggest their actual role includes a broad range of secondary roles, some of which are unanticipated. Thus, our review illuminates the breadth of roles CCC play in GME assessment systems. With a clearer understanding of the roles these committees *actually* play, programs are in a better position to advocate for resources to support CCC efforts. A more systematic approach to investigating CCC processes may not only assist CCCs in developing their expertise, but also provide meaningful input into how assessment systems can be improved to ensure fair and transparent assessment of learning as well as enhance assessment for learning (feedback and learning opportunities for trainees), contributing to achievement of overarching goals in CBME.

CCC performance ratings have been used as a source of validity evidence in the context of self-assessments, exploration of bias in assessment decisions, and prediction of performance in other settings—an

unanticipated role. This seems to suggest that CCC ratings are considered to be a "gold standard." Although findings from these studies may help critically analyze and improve assessment practices in specific settings, more work is needed to justify this approach. Further research into both anticipated and unanticipated secondary CCC roles can help training programs leverage the data they are already collecting to evaluate their practices and make improvements to their education and assessment systems. Therefore, we want to emphasize that even though secondary CCC roles (both anticipated and unanticipated) were not the original intent, they should not be minimized. More work is needed to determine the full scope of these roles and their effects on assessment systems, and we call for a collaborative national research agenda centered on studying specific CCC roles and their associated short- and long-term outcomes.

We also need to determine how implementation of CCCs contributes to achievement of intended outcomes in CBME (ie, graduating trainees who can provide safe and effective care). Although we found multiple examples of "CCC-inspired changes" or "outputs" such as new assessment tools, learning experiences, and data organization platforms, we did not find any studies that focused on investigation of long-term outcomes (eg, a change in feedback culture). To better understand if and how implementation of CCCs contributes to CBME goals, we therefore recommend that future studies consider using the logic model to study CCC impact, both short- and long-term, and intended as well as unintended outcomes.

There are a number of limitations to this work. Our review is limited to English-language studies only. Since studying the roles and associated outcomes of CCCs was not the explicit aim of most of the articles we reviewed, it is possible that we made inferences without fully understanding the contexts described. Although our taxonomy can serve as a first step for conceptualizing CCC processes and effects, it will need to be revisited and revised as future work is performed. Most of the studies were conducted in the United States and it is possible that our findings and conclusions may not be generalizable to other contexts. Thus, future studies should investigate CCC roles and associated outcomes in other contexts.

Conclusions

Although this narrative review identified broader CCC roles than anticipated, there were significant gaps in the literature regarding descriptions of these roles and their associated effects. The lack of articles specifically focused on investigating the outcomes of implementing CCCs in GME is a key finding and a launching point for future work.

References

1. Accreditation Council for Graduate Medical Education. ACGME Common Program Requirements (Residency). Accessed September 20, 2024. https://www.acgme.org/globalassets/pfassets/programrequirements/cprresidency_2023.pdf
2. Andolsek K, Padmore J, Hauer K, Ekpenyong A, Edgar L, Holmboe E. *Clinical Competency Committees: A Guidebook for Programs*. 3rd ed. Accreditation Council for Graduate Medical Education; 2020. https://www.acgme.org/Portals/0/ACGMEClinicalCompetencyCommitteeGuidebook.pdf
3. Sukhera J. Narrative reviews: flexible, rigorous, and practical. *J Grad Med Educ*. 2022;14(4):414-417. doi:10.4300/JGME-D-22-00480.1
4. Ekpenyong A, Edgar L, Wilkerson L, Holmboe ES. A multispecialty ethnographic study of clinical competency committees (CCCs). *Med Teach*. 2022;44(11):1228-1236. doi:10.1080/0142159X.2022.2072281
5. 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. doi:10.1111/medu.13884
6. 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. doi:10.1007/s40037-020-00624-x
7. Hauer KE, Chesluk B, Iobst W, et al. Reviewing residents' competence: a qualitative study of the role of clinical competency committees in performance assessment. *Acad Med*. 2015;90(8):1084-1092. doi:10.1097/ACM.0000000000000736
8. Mount CA, Short PA, Mount GR, Schofield CM. An end-of-year oral examination for internal medicine residents: an assessment tool for the clinical competency committee. *J Grad Med Educ*. 2014;6(3):551-554. doi:10.4300/JGME-D-13-00365.1
9. Beeson MS, Hamstra SJ, Barton MA, et al. Straight line scoring by clinical competency committees using emergency medicine milestones. *J Grad Med Educ*. 2017;9(6):716-720. doi:10.4300/JGME-D-17-00304.1
10. Schwartz A, Balmer DF, Borman-Shoap E, et al. Shared mental models among clinical competency committees in the context of time-variable, competency-based advancement to residency. *Acad Med*. 2020;95(suppl 11 Association of American Medical Colleges Learn Serve Lead: Proceedings of the 59th Annual Research in Medical Education Presentations):95-102. doi:10.1097/ACM.0000000000003638

11. Schumacher DJ, Bartlett KW, Elliott SP, et al. Milestone ratings and supervisory role categorizations swim together, but is the water muddy? *Acad Pediatr*. 2019;19(2):144-151. doi:10.1016/j.acap.2018.06.002
12. Schumacher DJ, West DC, Schwartz A, et al. Longitudinal assessment of resident performance using entrustable professional activities. *JAMA Netw Open*. 2020;3(1):e1919316. doi:10.1001/jamanetworkopen.2019.19316
13. Dehon E, Jones J, Puskarich M, Sandifer JP, Sikes K. Use of emergency medicine milestones as items on end-of-shift evaluations results in overestimates of residents' proficiency level. *J Grad Med Educ*. 2015;7(2):192-196. doi:10.4300/JGME-D-14-00438.1
14. Schumacher DJ, King B, Barnes MM, et al. Influence of clinical competency committee review process on summative resident assessment decisions. *J Grad Med Educ*. 2018;10(4):429-437. doi:10.4300/JGME-D-17-00762.1
15. Hamstra SJ, Yamazaki K, Barton MA, Santen SA, Beeson MS, Holmboe ES. A national study of longitudinal consistency in ACGME milestone ratings by clinical competency committees: exploring an aspect of validity in the assessment of residents' competence. *Acad Med*. 2019;94(10):1522-1531. doi:10.1097/ACM.0000000000002820
16. Kelleher M, Kinnear B, Sall DR, et al. Warnings in early narrative assessment that might predict performance in residency: signal from an internal medicine residency program. *Perspect Med Educ*. 2021;10(6):334-340. doi:10.1007/s40037-021-00681-w
17. Al-Bualy R, Ekpenyong A, Holmboe E. Effective clinical competency committee meeting practices: before, during, and after. *Acad Med*. 2023;98(11):1340. doi:10.1097/ACM.0000000000005123
18. Mamtani M, Scott KR, DeRoos FJ, Conlon LW. Assessing EM patient safety and quality improvement milestones using a novel debate format. *West J Emerg Med*. 2015;16(6):943-946. doi:10.5811/westjem.2015.9.27269
19. Cooney CM, Cooney DS, Bello RJ, Bojovic B, Redett RJ, Lifchez SD. Comprehensive observations of resident evolution: a novel method for assessing procedure-based residency training. *Plast Reconstr Surg*. 2016;137(2):673-678. doi:10.1097/01.prs.0000475797.69478.0e
20. Cullen MJ, Zhang C, Sackett PR, Thakker K, Young JQ. Can a situational judgment test identify trainees at risk of professionalism issues? A multi-institutional, prospective cohort study. *Acad Med*. 2022;97(10):1494-1503. doi:10.1097/ACM.0000000000004756
21. Rebel A, DiLorenzo A, Nguyen D, et al. Should objective structured clinical examinations assist the clinical competency committee in assigning anesthesiology milestones competency? *Anesth Analg*. 2019;129(1):226-234. doi:10.1213/ANE.0000000000004120
22. Bryant BH. Feasibility of an entrustable professional activity for pathology resident frozen section training. *Acad Pathol*. 2021;8:23742895211041757. doi:10.1177/23742895211041757
23. Rassbach CE, Blankenburg R. A novel pediatric residency coaching program: outcomes after one year. *Acad Med*. 2018;93(3):430-434. doi:10.1097/ACM.0000000000001825
24. Yen SPF, Lewis PJ. Implementation of a departmental resident web-based database. *Acad Radiol*. 2022;29(1):144-149. doi:10.1016/j.acra.2020.09.023
25. Stahl CC, Collins E, Jung SA, et al. Implementation of entrustable professional activities into a general surgery residency. *J Surg Educ*. 2020;77(4):739-748. doi:10.1016/j.jsurg.2020.01.012
26. Friedman KA, Raimo J, Spielmann K, Chaudhry S. Resident dashboards: helping your clinical competency committee visualize trainees' key performance indicators. *Med Educ Online*. 2016;21:29838. doi:10.3402/meo.v21.29838
27. Almufarrej F, O'Brien M, Shahait A, Nava G. Feasibility of smartphone application in plastic surgery operative assessments. *Plast Reconstr Surg Glob Open*. 2022;10(2):e4085. doi:10.1097/GOX.0000000000004085
28. Dougherty D, Weiler EJ, Sbraccia JC, Crowe CB, Lindsey HJ. A novel dashboard for improving the efficiency of a general surgery residency clinical competency committee meeting. *Am J Surg*. 2022;224(1 Pt B):363-365. doi:10.1016/j.amjsurg.2022.03.018
29. Johna S, Woodward B. Navigating the next accreditation system: a dashboard for the milestones. *Perm J*. 2015;19(4):61-63. doi:10.7812/TPP/15-041
30. Regan L, Cope L, Omron R, Bright L, Bayram JD. Do end-of-rotation and end-of-shift assessments inform clinical competency committees' (CCC) decisions? *West J Emerg Med*. 2018;19(1):121-127. doi:10.5811/westjem.2017.10.35290
31. Bedy SC, Goddard KB, Stilley JAW, Sampson CS. Use of emergency department pharmacists in emergency medicine resident milestone assessment. *West J Emerg Med*. 2019;20(2):357-362. doi:10.5811/westjem.2018.10.37958
32. Pack R, Lingard L, Watling C, Cristancho S. Beyond summative decision making: illuminating the broader roles of competence committees. *Med Educ*. 2020;54(6):517-527. doi:10.1111/medu.14072
33. Nabors C, Forman L, Peterson SJ, et al. Milestones: a rapid assessment method for the clinical competency committee. *Arch Med Sci*. 2017;13(1):201-209. doi:10.5114/aoms.2016.64045
34. Warburton KM, Goren E, Dine CJ. Comprehensive assessment of struggling learners referred to a graduate medical education remediation program. *J Grad Med Educ*. 2017;9(6):763-767. doi:10.4300/JGME-D-17-00175.1

35. Blau JA, Atia AN, Powers DB. Clinical competency committees in plastic surgery residency. *Plast Reconstr Surg Glob Open*. 2021;9(9):e3833. doi:10.1097/GOX.0000000000003833
36. Lyle B, Borgert AJ, Kallies KJ, Jarman BT. Do attending surgeons and residents see eye to eye? An evaluation of the Accreditation Council for Graduate Medical Education milestones in general surgery residency. *J Surg Educ*. 2016;73(6):e54-e58. doi:10.1016/j.jsurg.2016.07.004
37. Watson RS, Borgert AJ, O Heron CT, et al. A multicenter prospective comparison of the Accreditation Council for Graduate Medical Education milestones: clinical competency committee vs. resident self-assessment. *J Surg Educ*. 2017;74(6):e8-e14. doi:10.1016/j.jsurg.2017.06.009
38. Tichter AM, Mulcare MR, Carter WA. Interrater agreement of emergency medicine milestone levels: resident self-evaluation vs clinical competency committee consensus. *Am J Emerg Med*. 2016;34(8):1677-1679. doi:10.1016/j.ajem.2016.04.055
39. Ross FJ, Metro DG, Beaman ST, et al. A first look at the Accreditation Council for Graduate Medical Education anesthesiology milestones: implementation of self-evaluation in a large residency program. *J Clin Anesth*. 2016;32:17-24. doi:10.1016/j.jclinane.2015.12.026
40. Srikumaran D, Tian J, Ramulu P, et al. Ability of ophthalmology residents to self-assess their performance through established milestones. *J Surg Educ*. 2019;76(4):1076-1087. doi:10.1016/j.jsurg.2018.12.004
41. Meier AH, Gruessner A, Cooney RN. Using the ACGME milestones for resident self-evaluation and faculty engagement. *J Surg Educ*. 2016;73(6):e150-e157. doi:10.1016/j.jsurg.2016.09.001
42. Chow I, Nguyen VT, Losee JE, et al. Milestones in plastic surgery: attending assessment versus resident assessment. *Plast Reconstr Surg*. 2019;143(2):e425-e432. doi:10.1097/PRS.0000000000005214
43. Kwasny L, Shebrain S, Munene G, Sawyer R. Is there a gender bias in milestones evaluations in general surgery residency training? *Am J Surg*. 2021;221(3):505-508. doi:10.1016/j.amjsurg.2020.12.020
44. Santen SA, Yamazaki K, Holmboe ES, Yarris LM, Hamstra SJ. Comparison of male and female resident milestone assessments during emergency medicine residency training: a national study. *Acad Med*. 2020;95(2):263-268. doi:10.1097/ACM.0000000000002988
45. Zuckerbraun NS, Levasseur K, Kou M, et al. Gender differences among milestone assessments in a national sample of pediatric emergency medicine fellowship programs. *AEM Educ Train*. 2020;5(3):e10543. doi:10.1002/aet2.10543
46. Hauer KE, Jurich D, Vandergrift J, et al. Gender differences in milestone ratings and medical knowledge examination scores among internal medicine residents. *Acad Med*. 2021;96(6):876-884. doi:10.1097/ACM.0000000000004040
47. Athy S, Talmon G, Samson K, Martin K, Nelson K. Faculty versus resident self-assessment using pathology milestones: how aligned are we? *Acad Pathol*. 2021;8:23742895211060526. doi:10.1177/23742895211060526
48. Turner JA, Fitzsimons MG, Pardo MC Jr, et al. Effect of performance deficiencies on graduation and board certification rates: a 10-yr multicenter study of anesthesiology residents. *Anesthesiology*. 2016;125(1):221-229. doi:10.1097/ALN.0000000000001142
49. Kimbrough MK, Thrush CR, Barrett E, Bentley FR, Sexton KW. Are surgical milestone assessments predictive of in-training examination scores? *J Surg Educ*. 2018;75(1):29-32. doi:10.1016/j.jsurg.2017.06.021
50. Frey-Vogel AS, Scott-Vernaglia SE, Carter LP, Huang GC. Simulation for milestone assessment: use of a longitudinal curriculum for pediatric residents. *Simul Healthc*. 2016;11(4):286-292. doi:10.1097/SIH.0000000000000162
51. Schauer DP, Kinnear B, Kelleher M, Sall D, Schumacher DJ, Warm EJ. Developing the expected entrustment score: accounting for variation in resident assessment. *J Gen Intern Med*. 2022;37(14):3670-3675. doi:10.1007/s11606-022-07492-7
52. Roshan A, Wagner N, Acai A, et al. Comparing the quality of narrative comments by rotation setting. *J Surg Educ*. 2021;78(6):2070-2077. doi:10.1016/j.jsurg.2021.06.012
53. Donato AA, Alweis R, Wenderoth S. Design of a clinical competency committee to maximize formative feedback. *J Community Hosp Intern Med Perspect*. 2016;6(6):33533. doi:10.3402/jchimp.v6.33533
54. Abbott KL, George BC, Sandhu G, et al. Natural language processing to estimate clinical competency committee ratings. *J Surg Educ*. 2021;78(6):2046-2051. doi:10.1016/j.jsurg.2021.06.013
55. Rozenshtein A, Mullins ME, Deitte LA, et al. "What program directors think" II: results of the 2013 and 2014 annual surveys of the Association of Program Directors in Radiology. *Acad Radiol*. 2015;22(6):787-793. doi:10.1016/j.acra.2014.11.014
56. Lloyd RB, Park YS, Tekian A, Marvin R. Understanding assessment systems for clinical competency committee decisions: evidence from a multisite study of psychiatry residency training programs. *Acad Psychiatry*. 2020;44(6):734-740. doi:10.1007/s40596-019-01168-x
57. Sebesta EM, Cooper KL, Badalato GM. Program director perceptions of usefulness of the Accreditation Council for Graduate Medical Education milestones system for urology resident evaluation. *Urology*. 2019;124:28-32. doi:10.1016/j.urology.2018.10.042

58. Smit MP, de Hoog M, Brackel HJL, ten Cate O, Gemke RJBJ. A national process to enhance the validity of entrustment decisions for Dutch pediatric residents. *J Grad Med Educ*. 2019;11(suppl 4):158-164. doi:10.4300/JGME-D-18-01006
59. Duitsman ME, Fluit CRMG, van Alfen-van der Velden JAEM, et al. Design and evaluation of a clinical competency committee. *Perspect Med Educ*. 2019;8(1):1-8. doi:10.1007/s40037-018-0490-1
60. Ekpenyong A, Becker K. What resources do clinical competency committees (CCCs) require to do their work? A pilot study comparing CCCs across specialties. *Med Teach*. 2021;43(1):86-92. doi:10.1080/0142159X.2020.1817878
61. Ekpenyong A, Baker E, Harris I, et al. How do clinical competency committees use different sources of data to assess residents' performance on the internal medicine milestones? A mixed methods pilot study. *Med Teach*. 2017;39(10):1074-1083. doi:10.1080/0142159X.2017.1353070
62. Shumway NM, Dacus JJ, Lathrop KI, Hernandez EP, Miller M, Karnad AB. Use of milestones and development of entrustable professional activities in 2 hematology/oncology training programs. *J Grad Med Educ*. 2015;7(1):101-104. doi:10.4300/JGME-D-14-00283.1
63. Chen F, Arora H, Martinelli SM. Use of key performance indicators to improve milestone assessment in semi-annual clinical competency committee meetings. *J Educ Perioper Med*. 2017;19(4):e611.
64. Van Osdol AD, O'Heron CT, Jarman BT. "Same as it ever was"—only different: a successful ACGME innovative project to equip PGY I residents to take at-home call. *J Surg Educ*. 2014;71(6):e104-e110. doi:10.1016/j.jsurg.2014.06.002
65. Kearney AM, Rokni AM, Gosain AK. The Accreditation Council for Graduate Medical Education milestones in integrated plastic surgery programs: how competency-based assessment has been implemented. *Plast Reconstr Surg*. 2022;149(4):1001-1007. doi:10.1097/PRS.0000000000008938
66. Schumacher DJ, Martini A, Sobolewski B, et al. Use of resident-sensitive quality measure data in entrustment decision making: a qualitative study of clinical competency committee members at one pediatric residency. *Acad Med*. 2020;95(11):1726-1735. doi:10.1097/ACM.0000000000003435
67. Goyal N, Folt J, Jaskulka B, et al. Assessment methods and resource requirements for milestone reporting by an emergency medicine clinical competency committee. *Med Educ Online*. 2018;23(1):1538925. doi:10.1080/10872981.2018.1538925
68. Kou M, Baghdassarian A, Khanna K, et al. Guiding fellows to independent practice: current trends in pediatric emergency medicine fellow supervision. *Pediatr Emerg Care*. 2022;38(10):517-520. doi:10.1097/PEC.0000000000002676
69. Schumacher DJ, Poynter S, Burman N, et al. Justifications for discrepancies between competency committee and program director recommended resident supervisory roles. *Acad Pediatr*. 2019;19(5):561-565. doi:10.1016/j.acap.2018.12.003
70. Schumacher DJ, Martini A, Bartlett KW, et al. Key factors in clinical competency committee members' decisions regarding residents' readiness to serve as supervisors: a national study. *Acad Med*. 2019;94(2):251-258. doi:10.1097/ACM.0000000000002469
71. Schumacher DJ, Schwartz A, Zenel JA Jr, et al. Narrative performance level assignments at initial entrustment and graduation: integrating EPAs and milestones to improve learner assessment. *Acad Med*. 2020;95(11):1736-1744. doi:10.1097/ACM.0000000000003300
72. Stehman CR, Hochman S, Fernández-Frackelton M, et al. Professionalism milestones assessments used by emergency medicine residency programs: a cross-sectional survey. *West J Emerg Med*. 2019;21(1):152-159. doi:10.5811/westjem.2019.11.44456
73. Schumacher DJ, Martini A, Kinnear B, et al. Facilitators and inhibitors to assessing entrustable professional activities in pediatric residency. *Acad Pediatr*. 2021;21(4):735-741. doi:10.1016/j.acap.2020.11.013
74. Yaghmour NA, Poulin LJ, Bernabeo EC, et al. Stages of milestones implementation: a template analysis of 16 programs across 4 specialties [published correction appears in *J Grad Med Educ*. 2022;14(6):732. doi:10.4300/JGME-D-22-00841.1]. *J Grad Med Educ*. 2021;13(suppl 2):14-44. doi:10.4300/JGME-D-20-00900.1
75. Balach T, Conforti LN, Sangha S, Edgar L, Ames E. Qualitative study of orthopedic surgery milestones 1.0: burdens and benefits. *J Surg Educ*. 2022;79(5):1259-1269. doi:10.1016/j.jsurg.2022.05.002
76. Nabors C, Peterson SJ, Forman L, et al. Operationalizing the internal medicine milestones—an early status report. *J Grad Med Educ*. 2013;5(1):130-137. doi:10.4300/JGME-D-12-00130.1
77. Maranich AM, Hemmer PA, Uijtdehaage S, Battista A. ACGME milestones in the real world: a qualitative study exploring response process evidence. *J Grad Med Educ*. 2022;14(2):201-209. doi:10.4300/JGME-D-21-00546.1
78. Hanson MN, Pryor AD, Jeyarajah DR, et al. Implementation of entrustable professional activities into fellowship council accredited programs: a pilot project. *Surg Endosc*. 2023;37(4):3191-3200. doi:10.1007/s00464-022-09502-5
79. Lewis KO, Hathaway SB, Bratcher D, Blowey D, Knapp JF. Current milestones assessment practices, needs, and challenges of program directors: a collective case study in a pediatric hospital setting. *Cureus*. 2021;13(4):e14585. doi:10.7759/cureus.14585

80. Yamazaki K, Holmboe ES, Hamstra SJ. An empirical investigation into milestones factor structure using national data derived from clinical competency committees. *Acad Med*. 2022;97(4):569-576. doi:10.1097/ACM.0000000000004218
81. Schumacher DJ, Michelson C, Poynter S, et al. Thresholds and interpretations: how clinical competency committees identify pediatric residents with performance concerns. *Med Teach*. 2018;40(1):70-79. doi:10.1080/0142159X.2017.1394576
82. Conforti LN, Yaghmour NA, Hamstra SJ, et al. The effect and use of milestones in the assessment of neurological surgery residents and residency programs. *J Surg Educ*. 2018;75(1):147-155. doi:10.1016/j.jsurg.2017.06.001
83. Ketteler ER, Auyang ED, Beard KE, et al. Competency champions in the clinical competency committee: a successful strategy to implement milestone evaluations and competency coaching. *J Surg Educ*. 2014;71(1):36-38. doi:10.1016/j.jsurg.2013.09.012
84. Klutts JS, Guerin LA, Bruch LA, et al. Pathology milestones: assessing clinical competency by committee. *Acad Pathol*. 2015;2(4):2374289515614003. doi:10.1177/2374289515614003
85. Bienstock JL, Shivraj P, Yamazaki K, et al. Correlations between Accreditation Council for Graduate Medical Education obstetrics and gynecology milestones and American Board of Obstetrics and Gynecology qualifying examination scores: an initial validity study. *Am J Obstet Gynecol*. 2021;224(3):308.e1-308.e25. doi:10.1016/j.ajog.2020.10.029
86. Mink R, Herman BE, Carraccio C, et al. Agreement of program directors with clinical competency committees for fellow entrustment. *J Med Educ Curric Dev*. 2020;7:2382120520936613. doi:10.1177/2382120520936613
87. Mink RB, Schwartz A, Herman BE, et al. Validity of level of supervision scales for assessing pediatric fellows on the common pediatric subspecialty entrustable professional activities. *Acad Med*. 2018;93(2):283-291. doi:10.1097/ACM.0000000000001820
88. Mikhaeil-Demo Y, Holmboe E, Gerard EE, et al. Simulation-based assessments and graduating neurology residents' milestones: status epilepticus milestones. *J Grad Med Educ*. 2021;13(2):223-230. doi:10.4300/JGME-D-20-00832.1
89. Van Melle E, Hall A, Schumacher D, et al. Capturing outcomes of competency-based medical education: the call and the challenge. *Med Teach*. 2021;43(7):794-800. doi:10.1080/0142159X.2021.1925640

Andem Ekpenyong, MD, MHPE, is an Associate Professor, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois, USA; **Eric S. Holmboe, MD,** at the time of writing was Chief, Research, Milestones Development and Evaluation, Accreditation Council for Graduate Medical Education, Chicago, Illinois, USA, and is now Chief Executive Officer, Intealth, Philadelphia, Pennsylvania, USA; **Marjan Govaerts, PhD,** is an Associate Professor, School of Health Professions Education, Maastricht University, Maastricht, the Netherlands; and **Sylvia Heeneman, PhD,** is a Professor, Department of Pathology, School of Health Professions Education, Maastricht University, Maastricht, the Netherlands.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

Disclaimer: Andem Ekpenyong, MD, MHPE, works part-time for the Accreditation Council for Graduate Medical Education (ACGME), and at the time of writing, Eric S. Holmboe, MD, worked full-time for the ACGME.

The authors would like to thank Lisa Conforti, Raghdah Al-Bualy, Laura Edgar, and Lorelei Lingard.

Corresponding author: Andem Ekpenyong, MD, MHPE, Rush University Medical Center, Chicago, Illinois, USA, andem_ekpenyong@rush.edu

Received January 2, 2024; revisions received April 25, 2024, and August 21, 2024; accepted September 10, 2024.