2021-10-15

How do competence committees make decisions about resident progression? A qualitative study

Curtis, Colleen Mary

UNIVERSITY OF CALGARY

How do competence committees make decisions about resident progression?

A qualitative study

by

Colleen Mary Curtis

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN COMMUNITY HEALTH SCIENCES

CALGARY, ALBERTA

OCTOBER, 2021

Abstract

Competence committees (CC) determine trainees' progression through postgraduate competency-based medical education (CBME) programs. Models of how CC function identify that most programs take a problem-identification approach while others provide developmental feedback to every trainee. While CC are tasked with high stakes decisions, the process by which they discuss and make decisions about resident progression remains uncertain, with few publications addressing this question. The purpose of this qualitative study was to describe the factors affecting CC decision making. This instrumental case study examined two CC at a Canadian institution, three years post-CBME launch. Over a six-month period, one researcher observed four CC meetings and conducted interviews with 10 CC members which were audio recorded and transcribed verbatim. Royal College documents, CC terms of reference, investigator reflections, and memos created throughout the study were also examined. Following a constructivist grounded theory approach with constant comparison, two investigators coded transcripts independently and jointly to refine a codebook and identify themes in the data. Our theory-informed analysis led to a theoretical framework of CC decision-making: a process beginning within a social decision schema model and evolving to a discussion invoking social influence theory, shared mental models, and social judgment scheme to clarify the points of contention. The committee mindset determined the likelihood of entering a discussion about trainees; triggers for discussion related to CC members' uncertainty of the process or concerns with the adequacy of the data. Ensuing conversations considered the context of the individual resident and CC members' experiences. We found that ongoing challenges with CC functioning persist

three years post-CBME implementation. Despite Royal College recommendations and local terms of reference, CC provide limited developmental feedback to trainees who are doing well, and acknowledge that biases could affect the intended process. While this study only examined two CC, it identified important themes to address when considering a robust CC process.

Preface

This thesis is original unpublished independent work by the author, C. Curtis. The research reported herein was covered by ethics certificate REB19-0246, issued by the University of Calgary Conjoint Health Research Ethics Board for the project “How do competence committees make decisions about resident entrustment? A qualitative study” on June 13, 2019.

Acknowledgments

This work was made possible by the continual encouragement and support of my supervisor Dr. Lara Cooke. I appreciate the important feedback and insight provided by my advisory committee members, Dr. Aliya Kassam and Dr. Jason Lord. I also recognize the important contributions of the participants in this project who allowed me to learn from their perspectives of the postgraduate medical education system.

Above all, I thank Charles Curtis for his unfailing backing and inspiration.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

The implementation of competency-based medical education (CBME) represents a paradigm shift in postgraduate medical education to an outcomes-based curriculum (Frank, Snell, et al., 2010). In CBME, the determination of competence is envisioned to occur based on each individual trainee's demonstration of knowledge, skills and attitudes required for independent practice rather than time spent in training (Iobst et al., 2010). The determination of competence must be made by a purposively designed program of assessment within which all programs are expected to develop and utilize a competence committee (CC) to inform decisions related to resident progression and promotion (Andolsek et al., 2017; Holmboe et al., 2010). The CC are responsible for making high stakes recommendations based on their assessment of the level of trainee competence. As such, understanding the CC decision-making process and the factors that influence decisions ensures that residency programs are using a rigorous method to make their high stakes decisions. The aim of this study is to develop a theory to explain the processes used by competence committees to make decisions about resident competence and progression.

Chapter 2. Literature Review

Competence and Entrustment

CBME proposes a culture shift in the focus of training with establishment of predifined graduate abilities as the organizing principle for the program (Frank, Mungroo, et al., 2010). Competencies are observable abilities of the health professional and include multiple components such as knowledge, attitudes, skills and values; they can be assessed and assembled to inform resident development (Frank, Snell, et al., 2010). The competencies expected of a physician have been outlined in competency frameworks including the CanMEDS roles (Frank et al., 2015) and the Accreditation Council of Graduate Medical Education (ACGME) key competencies (Swing, 2007), among others. These provide a valuable starting point, however it is not practical to assess the competencies individually as they are intertwined in the daily practice of medicine. CBME describes a progressive attainment of competence, through scaffolding of expectations at different stages of training. The assessment of performance at each level, then, must consider the resident's performance on the competencies expected for their stage of training. To facilitate this conceptualization, ten Cate proposed the use of Entrustable Professional Activities (EPA) as the basis for assessment (2005). EPAs for a discipline are tasks that form part of the everyday work of a physician, require integration of the knowledge, skills and attitudes of one or more competencies, and can be observed and assessed in their performance. This approach has been adopted in practice, with the Royal College of Physicians and Surgeons of Canada (RCPSC) and ACGME adopting EPAs, competencies and milestones within their CMBE platforms as a way to guide progressive competency development (*Competence by Design cheat sheet*, 2016; Nasca

et al., 2012). Competence by Design (CBD) is the RCPSC platform of CBME, to be implemented by all specialty residency programs (*Competence by Design cheat sheet*, 2016).

Assessment in CBME

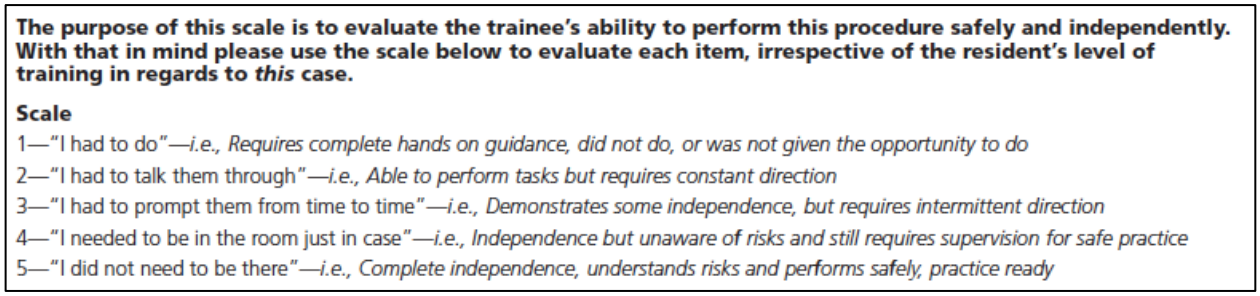
The determination of a resident's readiness for independent practice has long been comprised of completion of a time-based training program, local assessments of knowledge, skills and attitudes, and a culminating high-stakes certifying examination. The choice of local assessment methods were most often made based on convenience and accessibility of tools and examinations (Boulet & Durning, 2018). In the transition to CBME, postgraduate training programs have been challenged to adopt a purpose-built program of assessment that allows for competence to be measured and tracked over time. The generation of meaningful feedback to encourage progressive development is a key component of assessment in a CBME system (Holmboe et al., 2010). The strategy, then, is a system that includes frequent low-stakes assessments using different tools and different assessors, with immediate feedback to the learner, that can be combined to generate an overall picture of the trainee's competence and guide longitudinal coaching for improvement (Bok et al., 2013; Lockyer et al., 2017).

The validity of data collected is dependent on the instruments being appropriate for the task, accepted by the stakeholders, and well understood by the assessor (van der Vleuten et al., 2012). Entrustment has long been practiced in residency training as it concerns the supervising physician's assignment of tasks to the resident (ten Cate et al., 2016). As physicians are already in the practice of making entrustment decisions, this concept was incorporated in the assessment framework for CBD, in the form of an

entrustabilty scale that is intended to reflect resident competence for a specific task (Gofton et al., 2012).

The construct of entrustability was adopted for use in assessment with the development of the Ottawa surgical competency operating room evaluation (O-SCORE) (Gofton et al., 2012). This tool includes an intuitive scale that asks assessors to rate the amount of assistance required by the trainee for a specific task, shown in Figure 1. Trainees were deemed competent in their performance if they achieved “I needed to be there just in case” or “I did not need to be there”. As the acceptability of tools using an entrustability scale was reproducible (Rekman et al., 2016), this format was adopted by the RCPSC for use in workplace-based assessments (WBA) in CBD. Each specialty determined a list of EPAs that must be achieved by trainees to be deemed competent, and these activities are assessed in the moment using a WBA tool that includes the entrustability scale and a narrative comment (Gofton et al., 2017). Each WBA is a low-stakes assessment, reflecting only the single observation of the task described by the EPA. Trainees must collect many WBAs (along with other assessments determined by the individual programs) to provide sufficient information to support their competence.

Figure 1. *Ottawa Surgical Competency Operating Room Evaluation (O-SCORE) Entrustability Scale.*



The purpose of this scale is to evaluate the trainee's ability to perform this procedure safely and independently. With that in mind please use the scale below to evaluate each item, irrespective of the resident's level of training in regards to *this* case.

Scale

1—"I had to do"—*i.e., Requires complete hands on guidance, did not do, or was not given the opportunity to do*

2—"I had to talk them through"—*i.e., Able to perform tasks but requires constant direction*

3—"I had to prompt them from time to time"—*i.e., Demonstrates some independence, but requires intermittent direction*

4—"I needed to be in the room just in case"—*i.e., Independence but unaware of risks and still requires supervision for safe practice*

5—"I did not need to be there"—*i.e., Complete independence, understands risks and performs safely, practice ready*

Note: from Gofton et al., 2012

The intent of using EPAs and WBA is that these allow for direct observation and feedback of authentic encounters, which is thought to also provide more meaningful feedback to trainees (Holmboe et al., 2010). The performance of an EPA requires demonstration of multiple competencies that can be identified and targeted for improvement as appropriate. The utility of WBA and EPAs for assessment and to guide feedback in the moment has been confirmed, with the caveat that significant faculty development may be required for effective application of this tool (Tomiak et al., 2020).

While the implementation of CBD has required the use of EPAs as a form of assessment, it is recognized that a comprehensive program of assessment should contain data collected over time using a variety of tools and multiple assessors in order to optimize the quality and utility of the data collected (Schuwirth & Ash, 2013). Other sources of assessment data in a resident portfolio include the results of in-training examinations, observed structured clinical examinations (OSCEs), in-training evaluation reports, resident self-assessments, 360 evaluations, and others depending on the program. This combination of data should be reviewed in detail to provide a holistic view of the trainee, with different assessment tools targeting different aspects of physician development (Bok et al., 2013; *Competence committee guideline: process and procedures in decision making*, 2018).

Competence Committees

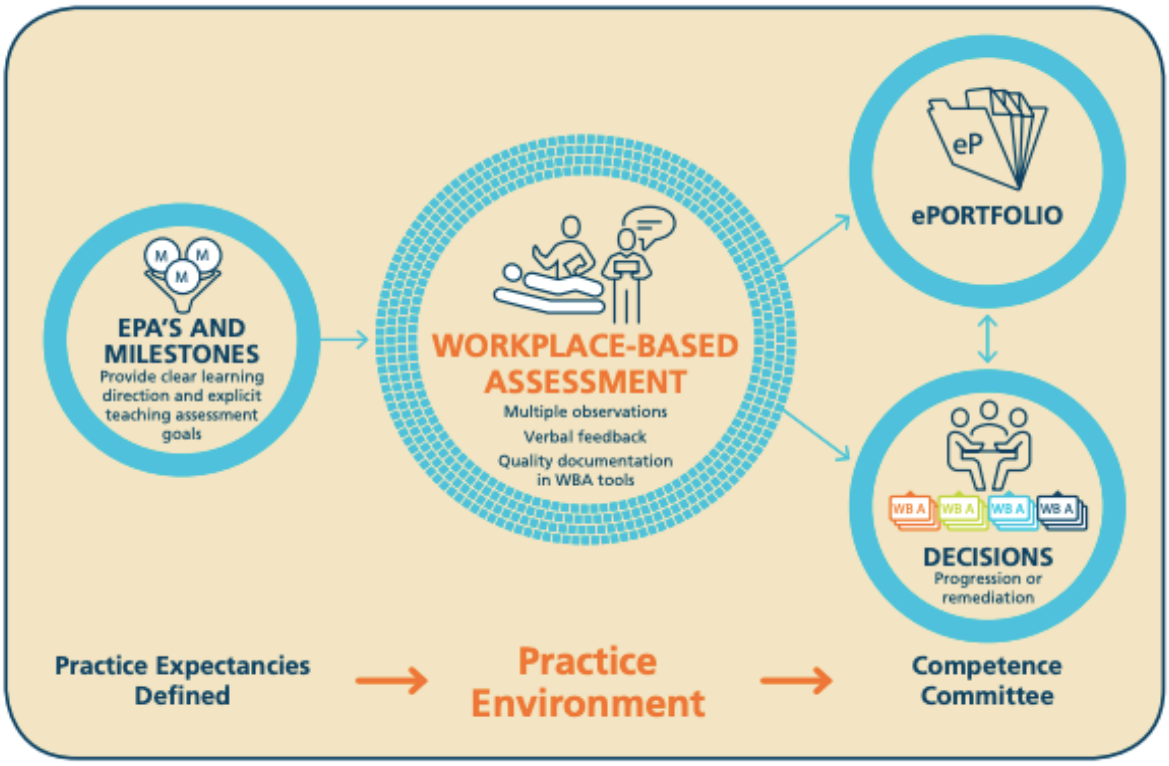
In the competency-based program of assessment, entrustability scores and feedback from frequent low stakes workplace-based observations are gathered in a resident portfolio along with other assessment results. The information is collated and reviewed through a process that can be effortful and requires that all contributing factors

are considered, including quantitative scores of performance and qualitative comments (van der Vleuten et al., 2012). The use of committees to review many individual assessments, based on different assessment tools, should generate useful feedback for the learner to help him or her progress, as well as making decisions about resident progression (Holmboe et al., 2010; Schuwirth & Ash, 2013); this is the basis for the CC. The fundamental purpose of the assessment program is that it should predict how a trainee will perform in the future and thus inform decisions about progression and the need for remediation (Duijn et al., 2018). This involves regular deliberations by CC, which are charged with the task of determining whether a resident is progressing at the expected pace and making decisions about promotion and entrustment of EPAs in CBD, as outlined in the RCPSC Competence Committee guideline (*Competence committee guideline: process and procedures in decision making*, 2018).

Canadian postgraduate specialty trainees progress through stages of CBD that are defined by achievement of successful EPAs and demonstration of competencies in other ways, without consideration of their post-graduate year of training. These stages are labeled *Transition to discipline*, *Foundations of discipline*, *Core of discipline*, and *Transition to practice* (*Competence by Design cheat sheet*, 2016). At every CC meeting, the status of the trainee must be set as one of: progressing as expected, progressing as expected with minor concerns, not progressing as expected, or progressing faster than expected. Along with this determination, the CC may promote a trainee to the next stage of training, recommend a period of remediation, allow them to sit a qualifying examination, or declare that training is complete and they are ready for independent practice (*Competence committee guideline: process and procedures in decision making*, 2018).

The flow of assessment information and decision making about resident status in CBD is demonstrated in Figure 2.

Figure 2. *Flow of Assessment Information in CBD Leading to CC Deliberation and Decision Making.*



Note. From Royal College of Physicians and Surgeons of Canada (Gofton et al., 2017).

The structure of CC in the United States was outlined by the ACGME (Andolsek et al., 2017) and the Canadian equivalent, by the RCPSC (*Competence committee guideline: process and procedures in decision making*, 2018); these minimum standards are designed to be adapted by programs as they see fit. Two paradigms have been described to explain the approach taken by CC early in the implementation of CBME: the problem identification model and the developmental model (Hauer et al., 2015). Although

resident assessment data is reviewed by CC members in both paradigms, how they examine this data and with what aim differ. In the problem-identification model, CC members use the assessment data to identify struggling residents, with the assumption that most residents will complete their training as competent physicians. This approach provides less meaningful feedback and support to the residents performing adequately, and in some cases identification of borderline residents may be delayed (Hauer et al., 2015). In the developmental model, CC use the data to help guide residents through the stepwise educational process towards mastery, believing that all residents will benefit from guidance to ensure progressive development (Hauer et al., 2015). This model of assessment and feedback draws from the basis of learning theory in medicine as proposed by Miller (1990) and the idea that progress up the pyramid requires active engagement of the learner (Lockyer et al., 2017). Using this developmental model for assessment of the learner's progress can maximize the probability that a graduating physician can provide care within the desired framework of competency. Notably, in Hauer's study, a majority of program directors followed a process most in keeping with a problem-identification model (2015). This approach to assessment and feedback is similar to the concept of mindset described by Carol Dweck who proposed that having a fixed mindset is a belief that your abilities are constant and unchangeable while "a growth mindset is based on the belief that your basic qualities are things you can cultivate through your efforts, your strategies and help from others," (Dweck, 2016, p. 7). Following from this, the description of the developmental model of CC function demonstrates a growth mindset, while a problem-identification model reflects a fixed mindset, at least with respect to the trainees without concerns (Hauer et al., 2015).

In recent descriptions of established CC, there is a consistent structure whereby the program director and other faculty members regularly meet to review learners' progress through defined EPAs and milestones (Chen et al., 2017; Ekpenyong et al., 2017; Klutts et al., 2015). The meetings usually include a focussed review of each learner with respect to each EPA, competency, or milestone, based on the individual assessments compiled (e.g. end-of -block assessment score and comments, structured portfolio, 360° evaluations, in-training examinations). Following the presentation, there is a discussion and judgement process on the adequacy of the developmental trajectory of the resident (Bok et al., 2013; Hall et al., 2018; Tekian et al., 2015). This suggests that program directors are trying to follow a process that aligns with the developmental model, where the overarching role of the CC is to determine the progress of a learner toward readiness for independent practice. However, the effectiveness of this process on the achievement of the stated goals remains uncertain. A deeper understanding of the decision-making process used by the CC could better support the judgments made on resident performance and result in improved implementation of CBME, by optimizing the assessment data collected, the synthesis of the data for review, and suggesting a structure for CC deliberation.

Decision Making

Small Group Decision Making

Summative entrustment decisions relating to trainee progression are proposed to standardize the level of supervision provided to individual residents. In this high stakes process, individual trainee factors related to the task should be considered, across all settings and supervisors; the validity of these judgements requires longitudinal

observations of the trainee across different contexts and by different supervisors (ten Cate et al., 2016). However, since these summative decisions are made by committee they also introduce dynamics relating to the group decision process. Multiple theories have been described relating to group decision making; while CC are relatively new, study of decision making in other fields has evolved over decades. Theories of decision making process most applicable to CC include social decision theory, constructivism, and social influence theory (Hauer et al., 2021).

First, social decision theory describes a structured method by which groups move from individual preferences to a group decision and emphasizes the importance of distinguishing between shared and unshared information (Chahine et al., 2017; Hauer et al., 2016; Kerr & Tindale, 2004). According to social decision theory, group members are more likely to believe and trust information that is shared amongst more members, and is consistent with their impression (Kerr & Tindale, 2004; Wittenbaum et al., 2004). This can lead to difficulty in accepting and contemplating novel information, although decision making is improved when more unshared information is considered and incomplete exchange of information can lead to poor decisions (Dennis, 1996). The challenge posed to the group is finding ways to encourage sharing and confidence in novel information; this can occur if the unshared information is held by a group member who is deemed to be particularly knowledgeable (van Ginkel & van Knippenberg, 2009). From social decision theory, it follows that making all group members aware that they each possess different knowledge and that they are to work on a task together that will benefit from integration of their ideas may lead to better decisions (van Ginkel & van Knippenberg, 2009).

Next, constructivism describes information processing as an active search for understanding to build a shared mental model (Chahine et al., 2017). Shared mental models are the representation of the group's common understanding of their task, interpretation of the environment and the required collaboration (Edgar et al., 2021). This is proposed to assist with decision making as the shared conceptual system of ideas allows the group to realize when a proposed solution is correct within that system (Kerr & Tindale, 2004). In order for the shared mental model to be effective, it must contain an accurate representation of the reality and be based on correct assumptions, there must be agreement between members about the group goal, and an understanding of how the group will work together to achieve these goals (Edgar et al., 2021). The application of the shared mental model proceeds by deliberating and processing information as it is introduced by the group members, until there is a shared understanding of the implications of the decisions (Chahine et al., 2017).

The composition of the group and relative input of each member also plays a role in committee functioning, this is the basis for social influence theory. This theory refers to the tendency for a group to be so strongly driven toward consensus that the members concede to a decision without thorough information processing (Chahine et al., 2017; Wittenbaum et al., 2004). This can occur because some influential group members convince others to change their impressions to align with the majority, or, as described in groupthink, because members want to maintain group cohesion and concede to the majority opinion, even if they have differing views. Three factors are described to promote *groupthink*: high cohesion of group members, insulation from outside ideas, and stress-inducing conditions, with the most important being the insulation of the group from outside

ideas (Wittenbaum et al., 2004). The need for diversity of opinion and independence of judgments is also expressed as a factor that moderates the ability of groups to make better decisions than individuals (Surowiecki, 2005).

Competence Committee Decision Making

The small group theories described above have begun to be applied to the study of CC functioning. In a CC that most closely follows a social decision schema, the main factors influencing faculty decisions regarding resident competence included the difficulty in interpreting comments and scores provided by a third party without context, and challenges relating to personal knowledge of the trainee in question (Donato et al., 2016; Ekpenyong et al., 2017). Importantly, Ekpenyong and colleagues (2017) concluded that the social decision schema was negatively impacted by the inability of the written comments to provide sufficient detail for effective sharing of knowledge. Similar uncertainty in how to interpret comments was described in a study of clerkship grading committees (Frank et al., 2019). These findings contrast with previous work on the utility of narrative comments on assessment forms (Hatala et al., 2017); perhaps it is an example of the social decision theory at play and the difficulty that cohesive groups have with the integration of unshared knowledge or, as suggested by Frank et al., that the committee members are not sure whether the assessors have the same shared mental model of the terminology used (2019).

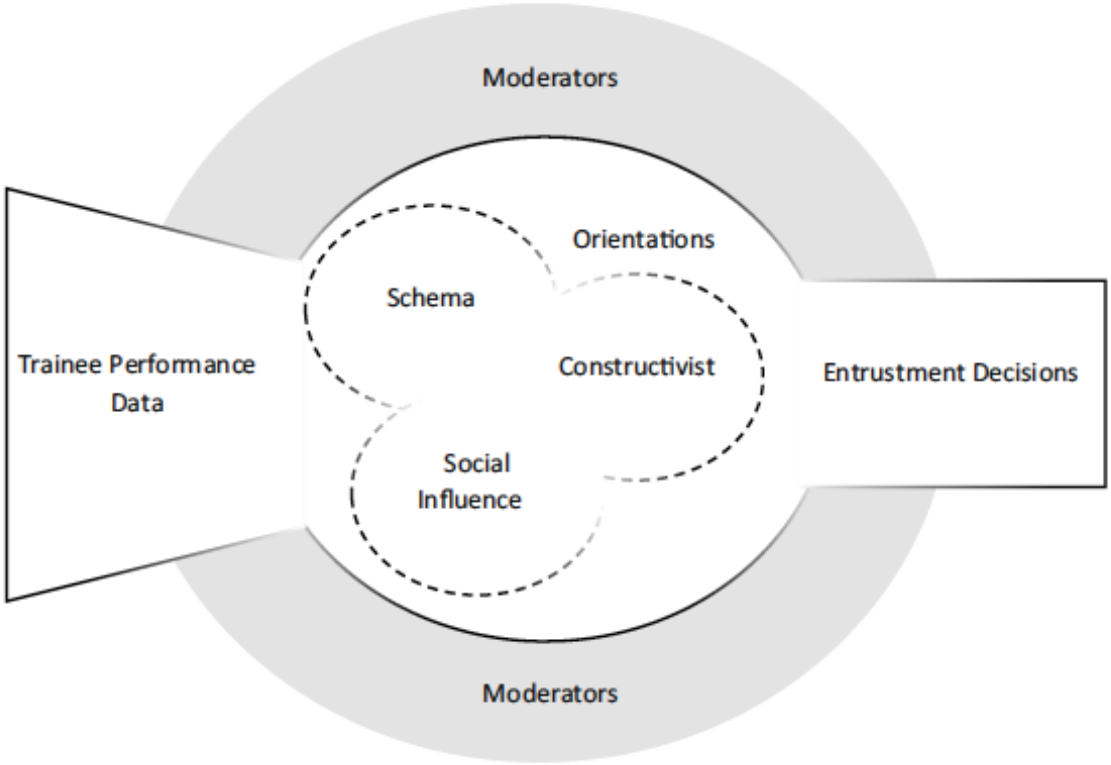
When combined, social decision schema and a constructivist approach may be more practical, as described in the evolution of the CC process at one institution (Chen et al., 2017). Chen related that the CC initially operated as a constructivist open discussion, then added a structured framework for the review of assessment information

and presentation to CCC members, which resulted in decreased meeting time and increased satisfaction without compromising the robustness of their decisions. A similar strategy was reported by Pack et al., who described that CC deliberations were schematic, linear, and straightforward unless problematic data is identified (2019). In those situations, CC members engaged in a constructivist discussion that involved close examination and unpacking of the data, debate, and attempts to verify findings.

The role of social influence theory in CC must be considered. It has been recognized that groups with high cohesion are more likely to justify their opinions and decisions rather than considering alternatives; most CC will have limited diversity, being composed of primarily academic/clinical faculty (Hauer et al., 2016). Although this is a concern, a study of a group review process of marginally performing clinical clerks found that one quarter of students received a different rating following group discussion (Gaglione et al., 2005). In that study, the main reason given for the changes in ratings was clarification of the information provided followed by group discussion, and more senior members of the committee were just as likely to change their rating as junior faculty. Similar results were described in group discussions of resident performance (Williams et al., 2005). These reports suggest that the dangers of groupthink can be mitigated by clear process, including ensuring all members are given the opportunity to speak, and encouraging junior members to voice their input first (Hemmer & Kelly, 2017). The robustness of CC decisions can be solidified by examination of the constructs used for decision making, including clarifying the role of groupthink in CC.

A unifying theoretical framework for CC decision making has been proposed by Chahine et al. (2017), encompassing the social decision, constructivist and social

Figure 3. *Theoretical Framework for Clinical Competency Committee Decision-Making.*



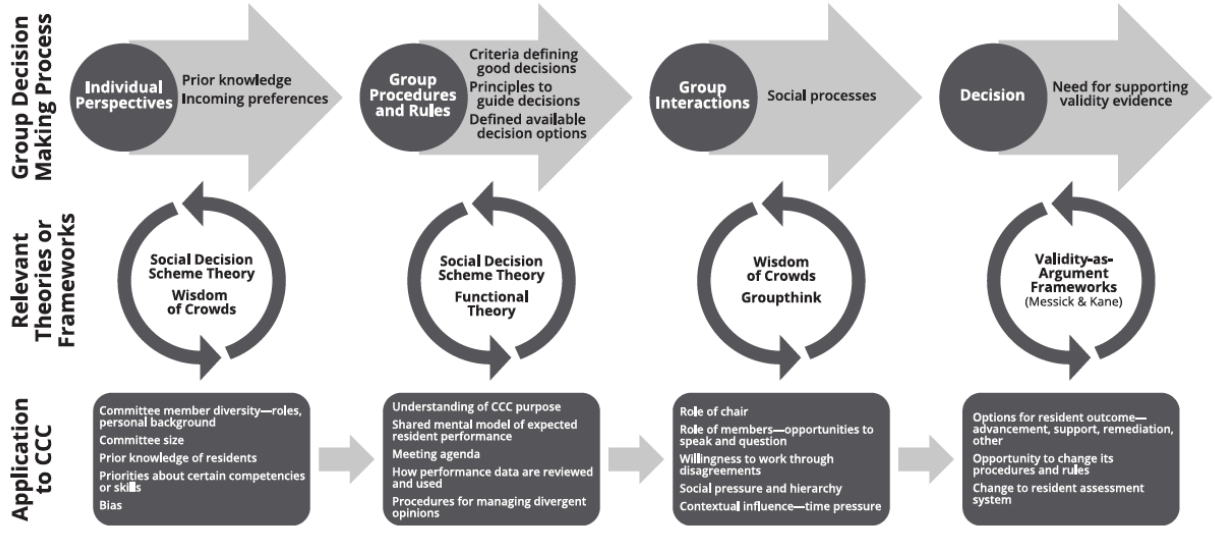
Note: From Chahine et al. (2017)

influence theories (Figure 3). This framework proposes that the CC process is centred on one or a combination of three orientations: schema, constructivist, and social influence. This orientation for decision making is moderated by factors such as guidelines, timeframes, and leadership, which affect the discussion of the CC about resident assessment data leading to a decision about resident performance and entrustability. The leadership, timeframes and guidelines within which the CC review assessment data and deliberate could encourage the committee to function within more of a problem identification model or developmental model. The relative importance of the proposed

moderating factors may also vary based on the program and the resident being discussed, further impacting decisions on resident progression as well as the quality of feedback that can be provided to the residents.

This model was refined by Hauer (2021) to describe that the CC process involves many steps which each invoke different theories of group function. As demonstrated in Figure 4, the four parts of the CC process considered are the individual member perspectives, group procedures and rules, group interactions, and decision outcomes. Similar to the theories already discussed, this model proposes that social decision schema, functional theory, groupthink and the Wisdom of Crowds can be combined to describe the process. They add Kane's and Messick's validity frameworks as a mechanism by which groups and their members consider the decisions and their consequences (Hauer et al., 2021).

Figure 4. *Group Decision-Making Framework Relevant for Clinical Competency Committees.*



Note: From (Hauer et al., 2021)

While these models form a starting point for understanding the CC decision making process, they are theoretical and largely did not include direct study of CC therefore how closely they reflect actual process is unknown. In order to ensure rigour of the process of decision making (to achieve the ultimate goal of safe patient care), examination of the CC decision-making process is warranted. This qualitative study aimed to explore the methods used by members of the CC of a residency program to make decisions about progression and competence of their residents. Specifically, the research questions were as follows:

1. How do members of the CC make decisions about the competence of residents?
2. What sources of data do CC members find most useful to inform their decisions?
3. How do CC members weigh presented data with their personal experience of the resident?
4. How does the composition of the CC affect the group decision making?

Chapter 3. Methods

Methodology

Theories have been proposed to explain the group decision making of CC, however it is unclear whether they truly explain the CC process. Our study aimed to develop a model that explained this decision-making process based on direct examination of operational competence committees. Case study methodology is ideally suited to explore and explain phenomena that are intrinsically linked with their contexts, including small group function and decision-making (Yin, 2018, p. 14).

Study Design

Our instrumental multiple-case study used in-depth examination of selected cases to describe and understand the phenomenon of interest, with cases selected for their ability to illustrate the phenomenon and inclusion of multiple cases to gain a deeper appreciation of the topic through replication (Crowe et al., 2011). The trainees discussed in CC meetings were at different stages of training and had different successes and challenges. Additionally, the members of each CC have varying seniority and experience with making entrustment decisions. The case study approach allowed us to disentangle anticipated complexities in the group relationships and interactions as we explored the factors affecting the decision process, to clarify how the competence and progress decisions were reached.

Case study research uses multiple sources of data to maximise the depth of information that can be collected from small samples, with the sample size depending on the richness of the data, variety of sources and phenomenon in question (Moser & Korstjens, 2018). The sample size required for grounded theory research is often

estimated at 20-25 interviews to reach thematic saturation (Guest et al., 2006). With the decision as the phenomenon of interest occurring multiple times in each meeting with different context, ensuring that rich data could be gleaned from each observation and interview. Therefore, combining multiple types of data including observation of meetings and data review, we estimated that 10-15 interviews would be required.

In order to protect the anonymity of the CC studied, this study has been reported in the form of a cross-case analysis with the emphasis on the phenomena of interest rather than the specifics of the cases. This methodology supported the internal and external validity of the research by considering multiple sources of evidence and existing theoretical frameworks to guide study design, data collection and analysis (Yin, 2018, p. 43).

Cases

The cases studied were the CC for two post-graduate training programs leading to fellowship with the Royal College of Physicians and Surgeons of Canada (RCPSC) at a single Canadian institution. The CC studied were purposively chosen for several reasons. First, they were well-established in CBD, having implemented the new style of training three years prior to the study. Second, they were different in size, being one mid-size and one small program with proportional difference in CC group size. Third, the researchers had no direct relationship with the CC or their trainees, minimizing any potential conflict of interest in the process of observation and analysis.

The bounds of the cases included all members of the CC during the time they spent reviewing files and in CC meetings as well as the operational documents in place through the study period. All members of both CC were invited to participate in the study

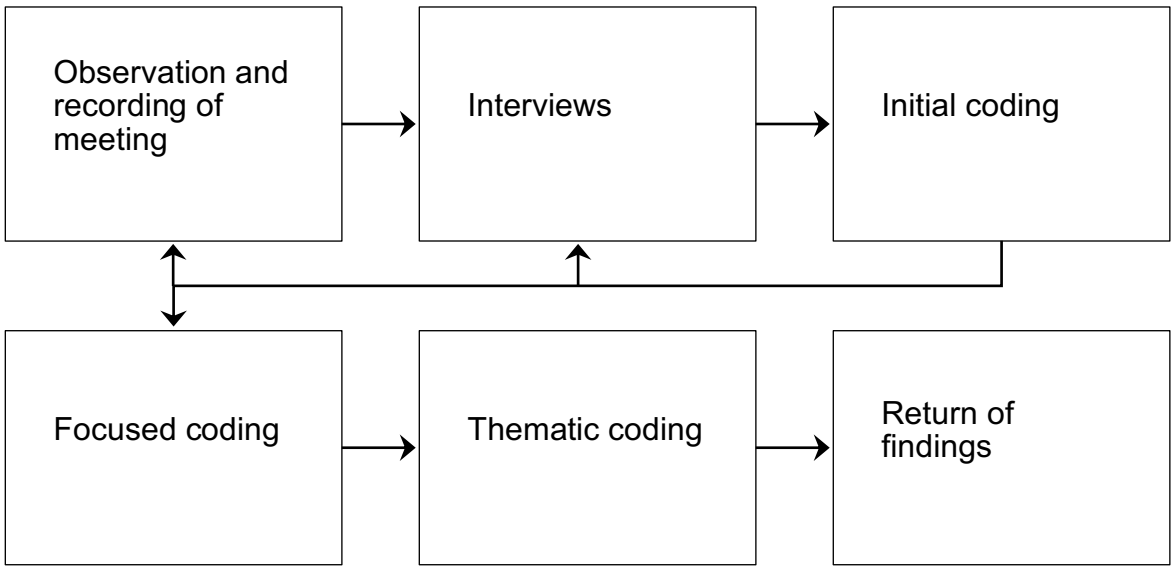
and provided written informed consent; the CC included in the study had full participation of their members. As the discussion of trainee performance is the essence of CC function and the subject of this study, trainees were also informed of the project and provided written informed consent for inclusion of discussions of their progress. Transcripts were anonymized with trainees referred to by a number and grouped by PGY-level, with no identifying information collected. The study was approved by the University of Calgary Conjoint Health Research Ethics Board (REB19-0246).

Data Collection

CC Meetings and Interviews

The study involved sequential meeting observation and one-on-one interviews (Figure 5). Over the course of six months, one researcher observed two meetings for each CC and kept field notes including observations and reflections. The same researcher conducted semi-structured interviews with 10 distinct CC members. Two or three interviews were held within the two weeks following each meeting involving six members of CC1 and four members of CC2. Due to the COVID-19 pandemic, all meetings

Figure 5. *Data Collection and Analysis Procedure.*



and interviews were conducted virtually using Zoom educational platform (Zoom Video Communications, Inc.). The meetings and interviews were audio recorded and transcribed verbatim for analysis (Transcription Heroes Transcription Services, Toronto, Ontario). Each CC (CC#), member (F#) and trainee (R#) was assigned a unique identifier and transcripts were de-identified prior to analysis.

The researcher contacted CC members directly for follow-up interviews to elaborate on observations from the meetings. The interview guide was established initially based on the research questions. The draft was pilot-tested with a program director not involved in the research project using a think-aloud strategy. Using this approach, the researcher asked the questions and the pilot respondent spoke out loud not the answer to the question, but rather how they heard and interpreted the question, voicing their thoughts out loud (Charters, 2003). The interview guide was then refined for clarity. Over time, the interview guide was modified as informed by iterative analysis, with questions added to clarify points of interest identified in early observations and interviews (Appendix A). Theoretical sampling was used to refine the number and choice of interviewees in the second and subsequent sets of interviews, based on the preliminary analyses of meeting observations and previous interviews. Specific members were invited for interviews if they were seen to have a key role in the discussions or a strong opinion that was contributing to the decision process. The researchers met throughout the study after each meeting and set of interviews to compare coding strategies, observations and thoughts. Data collection was halted after four meetings and 10 interviews, when the interim analysis demonstrated theoretical saturation. At this point the primary researcher presented the findings to the advisory committee who agreed that the data appeared comprehensive.

As there were no new ideas were emerging and the data contained ample depth to illustrate the variations within the themes (Charmaz, 2006, pp. 113-115).

Document Review

The researchers generated field notes, memos and annotations throughout the study; these were considered data and reviewed throughout the study process. We obtained discipline-specific EPA Guides, Requirements of Training, and Competencies documents from the RCPSC and terms of reference specific to the studied CC for inclusion in our analysis.

Data Analysis

Within the case study framework, a constructivist grounded theory approach to data analysis was employed. This approach, as described by Charmaz, recognized that the researchers brought experience with residency training, supervision and progression decision making and knowledge of intended CC process to their analysis, which contributed to the theory that emerged from the data (Charmaz, 2008). Data analysis began following the first meeting observed and proceeded iteratively throughout the duration of the study, using NVivo 12 (QSR International) for data management. Coding was done by primary investigator (C.C.) and research supervisor (L.C.) and proceeded in three stages: initial, focussed, and thematic coding (Thornberg & Charmaz, 2013). Initial coding was done independently as open coding with line-by-line review of the transcripts following the first CC meeting. The researchers met, reviewed the coding strategies and prepared a working codebook. After re-coding the initial transcript with the working codebook, the researchers again met to revise and refine the coding strategy until satisfactory agreement and understanding was achieved. Subsequently, the researchers

conducted independent and joint review of the transcripts, meeting frequently to compare consistency of coding and ensure that a mutual understanding of the data was emerging. Every 3 months, the researchers met with the thesis advisory committee, presenting the preliminary data review. This allowed for the addition of their reflections, questions and observations on the emerging analysis. With repeated review of the transcripts, researchers identified emerging themes, relationships and questions arising which were captured in annotations and memos. Focussed coding proceeded with categorizing and grouping of the initial codes to represent distinct themes. The codes and themes identified in the first case were interrogated in the second case according to strategies of replication logic and cross-case analysis (Yin, 2018, pp. 194-200). Particular attention was paid to similarities and differences between cases and specific decisions made during CC meetings as these were identified in the transcripts.

Thematic coding was structured within an organizing framework, demonstrating relationships between codes and themes that could be integrated to explain the committees' decision-making processes. The themes were refined by discussion with the research team and by referencing against the program documents. We identified illustrative quotes that were found to be most informative by demonstrating complementary or contrasting views of the theme. The preliminary findings and organizing framework were presented to participating CC members to allow for member checking and additional reflection by the participants.

Context

As qualitative researchers we were aware of the impact of our backgrounds and intervention on the process that we observed and interrogated. C. C. is a program director

in a postgraduate medical education specialty that is in the first year of CBD implementation and a graduate student in medical education. L. C. is a former program director and clinician educator with the RCPSC, involved in program evaluation. We considered our backgrounds and experience within a constructivist analysis as contributing to the framework, guiding observation and interpretation of the data. The thesis advisory committee shared their insights and reflections throughout the study conception and analysis; these individuals were post-graduate education researchers, knowledgeable with the intended goals and format of CBME. The return of findings sessions helped to ensure that our assertions reflect the data but also integrated the medical education knowledge and personal perspectives of the participants in the interpretation.

Chapter 4. Results

Data

The study included in depth assessments of two CC, situated in different residency training programs within the same Canadian institution. The transcripts from four CC meetings and 10 interviews, program-specific terms of reference, and RCPSC documents (source documents) together with the observations, memos and annotations generated by the study team formed the data for analysis. Specifically, the 10 interviewees included six members of CC1, three program directors or assistant program directors, two CC chairs and five members at large. A total of 28 trainee presentations and status decisions were observed over the course of the study, with some individuals presented more than once.

Case 1

The first case, CC1, was the competence committee for a five-year, medium sized residency training program leading to certification with the RCPSC. The program accepted four trainees per year who rotated through many clinical sites through their training. The CC membership included the program director, associate program director, and a trainee advocate as non-voting members in addition to the CC chair and six members of the clinical teaching faculty. One member was from a distinct but related discipline, three were actively involved in medical education research and two were former program directors. The program implemented CBD three years prior to the start of this study and the CC chair took on the role at the same time; the program had run a pilot CC prior to CBD implementation with a different chair. The program director had been in place for one and a half years prior to the study; they were previously the CBD lead for

the department. At the time of the study, two faculty had newly joined the CC while many had been members since its launch. Meetings were planned four times per year and lasted about three hours. Between eight and 16 trainees were reviewed at each meeting, with each individual reviewed at least twice yearly.

Trainees in this program were assigned faculty advisors who were not members of the CC but provided a narrative global assessment of how they viewed the trainee's performance to the designated data reviewer for the CC meeting. Reviewers were CC members who were tasked with reviewing all of the trainees' assessment data including previous CC reports, in-training evaluation reports, entrustable professional activity reports (EPAs), in-training examination results, feedback on teaching sessions and the global feedback from the faculty advisor. They presented a summary of this data at the CC meeting prior to proposing a motion on trainee status. If a seconder could be found for the motion, the committee then discussed as needed prior to voting on the motion. The CC determined the trainee status relative to their stage of training, but did not review in detail the EPAs attempted to determine whether each had been achieved. Rather, the PD would mark all EPAs for a stage as achieved once the stage of training was considered complete by the CC. The decision of the committee was returned to the trainee by the program director.

Case 2

The second case, CC2, was the competence committee for a small, two-year subspecialty training program leading to certification with the RCPSC. The program accepted one or two trainees per year and had a small group of core faculty members. The CC membership included all core faculty members, one was program director and

another served as chair; two members were former program directors. They were joined by one faculty member from a related program. The program director was the only non-voting member of the committee, and faculty advisors did not vote on their trainee's status. This program implemented CBD three years prior to the study and the chair had been in the role from the beginning. The program director had completed a graduate degree in medical education and had been in the role through the implementation of CBD, also serving as the section's CBD lead. Meetings lasted about 30 minutes and were scheduled 4 times per year. Two or three trainees were reviewed at each meeting, with each individual reviewed at least twice yearly.

Trainees in this program had faculty advisors who mentored them throughout the program, including representing them at the CC. The trainee was responsible for preparing a summary of their EPAs, research progress, teaching sessions, in-training examination results, study strategies and self-assessment. They reviewed this summary with their faculty advisor prior to the CC meeting; the faculty advisor also independently reviewed the original assessment data. The summary was presented at the CC meeting by the faculty advisor who proposed a motion on the trainee status. If a seconder was found for the motion, discussion was invited prior to voting on the motion. The CC determined the trainee status relative to their stage of training, but did not assess the completion of each EPA, unless this was relevant to the discussion of whether a stage was complete. The results of the CC meeting were returned to the trainee by the program director who scheduled one-hour meetings with each trainee following the CC.

Coding

The initial coding structure included codes relating to the role and process of the committee, the data used by the committee and the group dynamics and discussion. Through repeated review and analysis of the data, we identified the prominent themes as relating to process, data, discussion triggers and interpretation; the initial codes were reorganized within the new structure as demonstrated in

Table **1**. The “Mindset” theme was identified during coding, situated within the CC role, process and interpretation codes; it was separated out later as we recognized the intrinsic importance of this factor in the CC discussions. The themes are explored in detail in the following sections.

Table 1. *Initial Codebook Reorganized to Themes*

Initial Coding	Thematic coding
CC role	Process
CC evolution over time	process uncertainty
Data	CC evolution over time
narrative vs numbers	Data
personal experience	narrative vs. numbers
relativism	data sufficiency
trainee context	approach to EPA collecting
Discussion triggers	Discussion triggers
Group dynamic	Interpretation
Interpretation	personal experience
Process (uncertainty)	effects of discussion
Software	trainee context
	Mindset

Theme 1: Process

Process Uncertainty

The joint understanding by members of the role of the CC and the process that it should follow was fundamental to the decision-making process. For trainees who were doing well and more so for those who were struggling, questions about process arose. “I think there remains amongst the committee members some level of uncertainty with regards to by the book EPA counts that are needed to progress a trainee, versus overall gestalt based on their performance.” (interview, F19). In multiple situations, the committees were unsure what criteria they should consider for trainee progression between stages. As one example, stated during CC1 meeting,

> Are the deficiencies we've noticed significant enough that you would want to hold them back out of core? Do you want everything to be 4s and 5s before you progress them? Or are we saying, “Yeah, there's a couple of deficiencies that maybe they should progress with those noted and we put something together to help them moving forward”. I'm not quite sure what the criteria is. (CC1 meeting, F6)

The implications of designating a trainee as progressing faster than expected and advancing early were discussed at length in one meeting, illustrating the difficulty in transitioning between time-based and competence-based training. This idea was clarified in interviews, with one member stating:

> I think that's something that was not really thought about very carefully when EPAs were introduced - the question about the time frame. If you have

> someone that finishes everything within half the time, where does that put them, can they graduate sooner? Maybe not? (interview, F13)

The importance of process uncertainty also included the need to be reminded of the goals of the committee. CC members recognized difficulty in keeping their discussion to the data in the file, and the need to return to basics to exclude personal views of trainee competence.

> I think one of the things that caused the most [discussion] was not necessarily the views of the trainee. I found they were actually relatively homogenous. It was more what is our job as a CC; if our job is to consider feedback as objectively as we can based on what is provided to us, that is fundamentally different than our personal opinions of how well a trainee is doing. (interview, F8)

Evolution over Time

Members of both CC identified that the process and functioning of their group had developed over time. Their experience with the CC process made the CC more comfortable in their roles and trusting of the process. This was described by one member as:

> Traditionally, this committee reacted later than it ought to have in some cases. Now, the committee is more willing to make a decision [to identify a trainee as not progressing as expected] earlier on and trust the information and trust the process. Now that we've made that decision about multiple trainees who've gone on to successfully get through, people are more trusting of the process that's outside of this committee. (Interview, F5)

In addition, the experiences of the CC enabled them to adapt their structure to balance efficiency and completeness of discussion. Both cases used a template for presentation of data, that they felt ensured a thorough review of each trainee:

> Using the progress checklist has really helped us be more holistic in speaking about the trainees. We've had it since the beginning, but now that we have more trainees who are in CBD, it does [organize] the discussion for each trainee. (interview, F12)

With more experience, the CC were able to recognize that a lack of faculty development within the assessor pool and for CC members made it difficult to understand and interpret some EPA data. The complexity of the EPA assessment process and understanding what each was designed to measure was identified as an evolving challenge:

> The EPAs have multi-part subsections. This has been a learning curve in that when we're evaluating our trainees, we're not diving deep enough to look at each subsection to figure out whether they accomplished all of them. We've tended to write off an EPA as being completed when parts of it may not have been. (interview, F15)

Theme 2. Data

Narrative vs. numbers

High quality assessment data was identified as essential to the process by the CC members. Many problems with data were identified, most commonly the need for illustrative and specific comments from observers. This was especially important when there was discrepancy between the entrustability score and the narrative comment:

> The EPA may have been scored as a 4 or 5 but the comments did not match that. Or the EPA was scored as 3 but the comments were “they were totally independent”. […] the Competence Committee is really only able to make global decisions when the data that they’re given is accurate and understandable. (interview, F9)

The CC members reported using the comments to help identify flags and verify that the entrustability score assigned accurately reflected the performance. This related directly to a second concern identified with EPA assessment scores: that some observers did not seem to understand the goal of the EPA or consider the context in which they were conducting the assessment. This was reflected in multiple comments, eg. “The process has become a lot easier as preceptors become more acquainted with CBME, and have started to give more targeted feedback and understand the EPAs a bit better” (interview, F9), and “Many of them will have gotten low scores on that EPA but it's actually a misunderstanding of what the EPA is intended to measure”, (interview, F2).

Data Sufficiency

The CC members were clear that the data contained in EPA assessment forms alone was insufficient to identify trainees who were not progressing as expected. The EPA data reflected assessments of select observations, most often requested by the trainee. This was described by one member as:

> We struggled for the first few years that nobody had any of the lower scores, they were all 4 or 5 and there was a sense that they wouldn't ask [for an assessment unless the trainee felt they were fully entrustable]. The trainees would wait until the shift was done and upon hearing “you did a good job on

> that [case]" they would ask, "can you fill in that EPA for me". You knew that they were only targeting successful EPAs. (interview, F2)

This led to a suspected over-representation of successful observations. Members also questioned whether the absence of low-scoring EPAs should be considered as a point of concern, "I actually worry more about the ones that are consistently getting only fours or fives, that they're not putting themselves out there on the more challenging cases", (interview, F2).

In addition, not all aspects of medical practice were contained in the EPAs for a discipline. There were knowledge and performance metrics outside of the EPAs that merited consideration such as examination scores, "There have definitely been trainees who are progressing fine through EPAs but are having examination difficulties" (interview, F5), and professionalism, "We had a trainee that was not progressing as expected, they had not completed the number of EPA observations in the expected amount of time, but more so was having professionalism concerns, with answering emails or completing documents or consults", (interview, F12).

Approach to EPA Collecting

The global evaluation of EPA assessments by the committee was further complicated by trainee approach to EPA completion and the CC members' consideration of EPA count. Both CC expressed that simply achieving the minimum count of EPA was insufficient evidence to judge that a trainee had achieved competence for that task:

> It's very tempting just to count the number of EPAs and say it's good enough. But the minutiae of the EPAs, it doesn't really allow that. And it's

> always quite clear that the number of EPAs really doesn't matter as much as how the person is actually performing." (interview, F16)

Some trainees were recognized to be motivated to collect EPA observations and therefore achieved the minimum counts quickly, "One was able to fulfill [many more EPAs than their peers] within the same time frame of training. They were much more efficient in identifying what might apply as an EPA and sending lots of requests", (interview, F13). Conversely, the failure to collect sufficient EPA observations could identify real problems in trainees, "trainees that struggle seem to be "less good" at getting EPAs done. If they are barely meeting the minimums it's usually a hint that there is a problem - but it may reflect problems with executive functioning" (interview, F2).

There was minimal attempt by either committee to determine their trainees' level of entrustability for any individual EPA, but rather they relied on the primary reviewer's opinion of the count of achieved EPAs. The reasoning for this was explained by one member as:

> Our program is just not EPA driven. We tell them they need to get their EPAs, I would say our learners target the bare minimum because there's only so much value you can get out of collecting EPA's. […] I just don't think you can break the specialty down into these discrete bundles of presentations where you either achieved what somebody thought was appropriate for independent practice or not. (interview, F2)

Theme 3. Discussion Triggers

The CC meetings progressed in a similarly structured way in both cases studied. The opportunity for discussion was presented for every trainee after the presenting faculty

member put forward a motion on their status. However, discussion could also be triggered by a single question or dissenting comment and aimed to clarify the presence and source of a problem. The question could arise based on the data presented to the committee, “the reviewer proposed progressing as expected […] but people heard comments coming through the objective feedback that have made them think otherwise. Then we've had more in-depth discussion because it's been flagged by someone who's seen patterns arising” (interview, F8). It could be implied by the reviewer that there were concerns that warranted discussion, but most often it was another member who triggered the discussion:

> When file reviewers have some small concerns, I think it’s easy when a trainee is early in their training to push those concerns aside and say, ‘OK, will I go with this for the time being?’ As soon as one person voiced that maybe there was an alternate path to be considered, or a couple of higher-level concerns, it did give everybody else permission to voice their concerns and then consequently come to another decision. (Interview, F9)

While a single comment in the data presented could raise questions, there was cognizance amongst CC members that there should not be too much weight placed on any one observation. The ensuing discussions tried to clarify how important the comment was, and whether it required action:

> I worry sometimes that the comments are actually overvalued. We were talking about one trainee whose basic progression all looked fine, except for one comment. And there was a significant amount of discussion about that one comment. I think that’s OK, in the sense of the committee being

informed. But I was concerned at the time, that we were going to put too much weight on this comment. (Interview, F5)

Theme 4. Interpretation

Personal Experience

While the goal of the CC to review documented assessments and make an objective, fact-based decision was clearly described in both CC terms of reference, the committees found it difficult to separate their personal views of the trainee from the data. The members recognized the need to separate personal anecdotes from the data in the file, however it was more difficult to ignore the impression that members had formed in working with a trainee over years. One member was able to identify this as a form of bias, but the concept was generally described as the members having personal, implicit impressions of a trainee's competence. Illustrative quotes for this theme are presented in Table 2.

Table 2. *Quotes Describing the Effects of Personal Experiences with Trainees on Data Interpretation*

Committee members do bring in their biases; I think that those biases are informed by a tacit judgment. You can't build a rubric that says 'is this person at their level or not' […] when you don't have a comparator and you're saying "how is this person doing on these competencies?" It's incredibly difficult. (Interview, F2)
You try to be as objective as you possibly can when reviewing the chart but your own preconceived notions can't help but be involved in what you're presenting. (interview, F6)
if you've worked a lot with a trainee, they are now not just numbers on a page and words on feedback. You have a personalized vision of what that trainee's performance is like. (interview, F10)

Effect of Discussion

Despite their preformed impressions of trainees, CC members tried to listen to the data presented and participated in discussion to make decisions about trainees. They described that the discussion rarely changed their opinion about the trainees' competence, but more often provided evidence to clarify their progress decisions. This was described by one member as:

> It was through the discussions and what the Program Director and Assistant Program Director were adding from their perspective and with the EPAs that I was able to say, 'I feel really comfortable with the idea that this trainee needs to be reviewed sooner.' […] I think the group discussion was what really helped to solidify sort of the correct decision for that trainee taking all of these different factors into consideration. (interview, F6)

In some circumstances, members reflected that although the group discussion changed the progress decision from the initial motion, there was little difference in their impression of the trainees' competence. "No [the change in progress status] didn't change [how I think about the trainee], it was really nuance to give the right message to the trainee […] not to change our general impression about the trainee," (interview, F13).

The members acknowledged that the program directors were often able to provide context or explanations that were beneficial in their interpretation. In the discussion, there were requests for program directors to add information such as:

> The thing this trainee is working on that's an ongoing minor concern would be that they have a lot of comments about efficiency. And I'm not sure

> whether maybe [program director] are able to comment on this. I didn't see a note of how many patients they're seeing, but universally the evaluations do say that they're working very hard at improving efficiency, taking a very stepwise approach to this. (CC1 mtg, F11)

The discussions were felt to be balanced with no overriding dominant voice and questions could be raised comfortably by all members. Despite having a wealth of information about the trainees, the program directors were respectful of the intended process:

> The program directors and assistant program directors, […] have mainly provided context and rarely provided thoughts as to where the level ought to be. […] I mean a risk would be that they would be a dominant voice and they've not been. They've been an informative voice. (interview, F5)

Trainee Context

Data interpretation was affected not only by the faculty's personal experiences with the trainees and the assessment data, but also the trainee's context and history. CC members acknowledged that trainees with a history of difficulties may have gotten a closer review of their file, "Once there is even a little flag or minor concern noted, that tends to roll forward with the trainee for at least a year or two", (interview, F8). This could have led to a changed expectation for the trainee based on past performance, despite the stated CBME goal of norm-referenced assessment. This feeling was expressed with respect to a borderline trainee, "This trainee, every single review, is always the same. They're progressing as expected for them, perhaps not compared to all the others that

we are assessing, so for me, “progressing as expected for the candidate” fits”, (CC2 meeting, F18).

Trainee context could also have affected the approach to progression decisions when there were suspected to be external factors affecting performance. For some trainees, their response to prior feedback made the CC hesitate in making a determination:

> There was a significant conversation about how that decision [to not progress to core of discipline] would impact the trainee’s mental health. In my view, that’s not the role of the committee. If the trainee is not progressing as expected, they deserve to know immediately so that they can have support. Delaying that because people have a big heart and are worried about the trainee, I don’t think it’s actually in the trainee’s interest to do that. (interview, F5)

The presence of major external factors affecting performance was generally not identified with single discrepant assessments, but rather reflected by the entirety of a portfolio, and may have required additional action beyond the CC scope:

> The bigger challenge is when you see those multiple unsuccessful attempts that are global and pervasive. Then there's a bigger problem going on and I'm not sure you can always tease over what the problem is just by looking at the evaluations. […] sometimes it is just an area of knowledge, sometimes it’s a global life problem and there's a giant issue that you need to deal with on different levels. (interview, F10).

Theme 5. Mindset

A common theme recurring in the data was the mindset of the CC. Some individual members did demonstrate a growth mindset in their contemplation of trainee progress. The trajectory of entrustability scores was considered as a way to visualize progress:

> If the EPA scores look like a large percentage of them are in progress [scoring less than 4-5], then I'll look at the trend. If it looks like there's an overall trend that shows they've been doing better over the course later in the EPA, then great. (Interview, F8)

Another CC member described the process of feedback for learning as a key component of CBME and described that on one occasion, “the trainee said ‘I know that I'm not going to pass this’ and still wanted it evaluated. I appreciate that. […] filling out the EPA is not as important as sitting down and giving the trainee feedback on the performance”, (interview, F15).

In contrast, some members consistently described and demonstrated a fixed mindset with respect to the CC process. The difference in the time spent discussing and reviewing strong trainees compared to their peers was noticeable, along with the lack of recommendations provided for them,

> As far as the trainees who are doing well goes, there’s always the risk that those trainees fly by and are not necessarily either pushed or offered the level of constructive feedback that they should be getting. But whether or not the Competence Committee is the best place to identify that I’m not certain of. That seems to be more of an issue for the faculty advisors or Program Directors to assist with.” (interview, F9).

This was further clarified by two members' responses to the question "*Would you consider rating a trainee as "progressing faster than expected"?":*

> I don't think I'll be pursuing having a trainee graduate early simply because there's so many required training experiences that you can't get through in an appropriate fashion while still doing all of your rotations. And getting yearly opportunities to do their med-ed day which is not the same experience every time, they sort of build through the program in those experiences so I think they would miss out substantially on some of those experiences if they were rushing through. (interview, F2)

> We have never made the decision that someone's progressing faster than expected. And that may not be true, right. Among the trainees, there's one who is clearly progressing faster than the other trainees. Like they are objectively progressing faster, but there's definitely a reluctance to make that as a determination. (interview, F5)

CC members described that they felt the role of the CC was to identify whether trainees were progressing as expected or not, and that further determination around goals and growth should take place in discussion with the program director or faculty advisor,

> I think our Committee is much more focused on the initial sort of decisions and progress; we'll make the decision that they need to be remediated or they need to have a learning plan. But that then falls on the Program Director to come up with what the learning plan and objectives are going to be. Then it's our job to come back and say 'yes they've met them' or 'no they haven't'. (interview, F6)

> I see this committee as a bit of a screening for trouble and for identification of trainees who are having trouble and the like. […] this seems to be a lot more about the process of ensuring the residents are on track and doing okay and the work with the program directors and sort of our education hubs colleague is really where pushing new experiences is happening. (interview, F10)

This understanding of the roles of the CC and program director was in line with the CC1 terms of reference that described the role of the CC as advisory to the PD:

> A regular review of trainees' progress facilitates a developmental approach, supporting trainee learning over time. The CC should help the education team identify trainees who are not meeting their milestones and can suggest or mandate support and coaching for the trainee before the trainee gets too far off their trajectory. (CC1 terms of reference; Appendix B)

While the CC2 members demonstrated a similar approach to process and feedback to the trainees, the CC2 terms of reference outlined the role of the CC without reference to the program director:

> The mandate of the Competence Committee is to review and discuss learner portfolios in order to:
>
> - Advise and guide trainee learning and development;
> - Adjust a trainee's training experiences to enhance learning opportunities;
> - Review assessments to determine a trainee's achievement of each Entrustable Professional Activity (EPA);

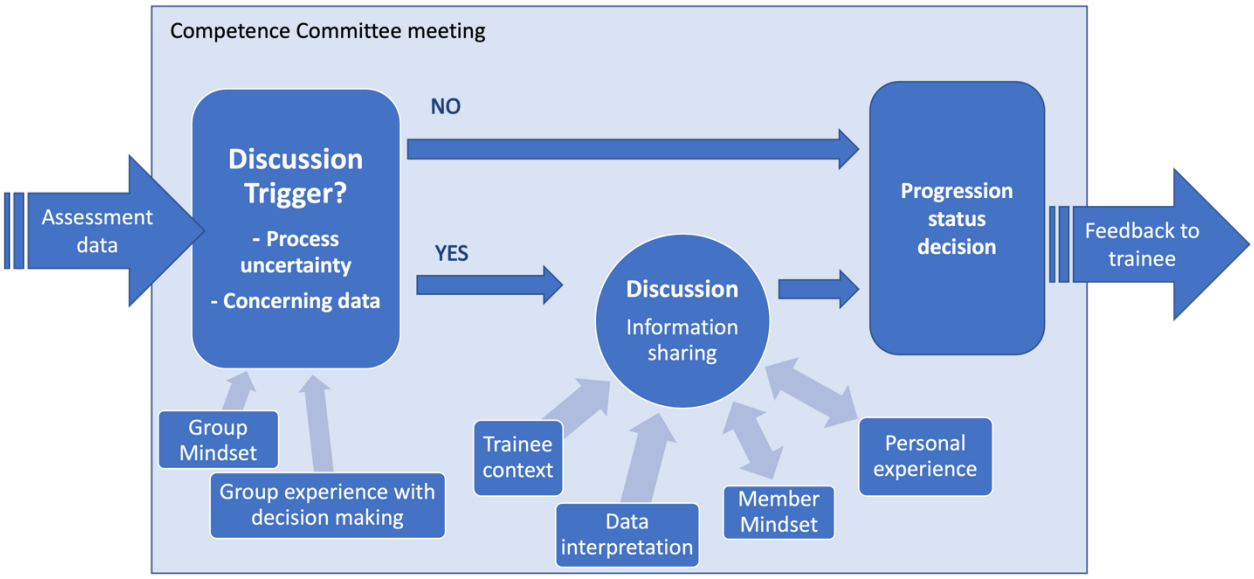
- Recommend learner status changes and progression of trainees through stages of training based on achievement of EPAs. (CC2 terms of reference; Appendix C)

Committee Process Model

An understanding of the CC decision-making process emerged from our analysis and the resulting model is shown in Figure 6.

The CC process followed primarily a social decision scheme that was moderated by the committee mindset, trainee context and experience of the group and individual members. Information flowed into the committee meeting as reviewed and presented by a CC member. The proposal and seconding of a motion could then proceed toward a decision one of 2 ways. The initial decision was made as a determination of whether the trainee was clearly meeting expectations or not. If the trainee was perceived to be doing well and the mindset of the committee was not growth-oriented, no discussion ensued. The identification of uncertainty regarding trainee status was often triggered by a single comment by a CC member who perceived a problem with the data presented or the process to be followed. In these situations, discussion ensued as members shared additional information to clarify the problem identified and come to a shared understanding of the situation. There was a deliberate attempt to identify unshared information that could inform the other group members. The depth of discussion was affected by the individual trainee's context, the prior experiences of the CC members individually and the group experience with decision-making. After the problem had been clarified to the satisfaction of the members, a decision was made about the trainee's level of competence and progression.

Figure 6. *Model of the Competence Committee Decision Making Process*



The thorough analysis of a discrepant example wherein a high-achieving trainee's status decision did trigger intense debate provided additional insight into the CC members' views. As examination of discrepant cases is often informative to explain a process, this discussion was reviewed specifically with a lens to generating questions for subsequent interviews and informing analysis of other meetings. In the case of the high-achieving trainee, the discussion was triggered by a motion that they were progressing faster than expected and should be promoted to the next stage of training early, ahead of their cohort. The ensuing conversation included a review of the process expected of the CC, questions about the quality of the completed EPA assessments, and systemic questions about the role of time in the training of CBD trainees. At no point were suggestions of how to encourage ongoing progression of skills and development of excellence raised, which paralleled other discussions of high achieving trainees. The triggers and themes identified during the discussion were the same as for

the trainees who were flagged as potentially struggling; clarification of the process and requirements to progress trainees between stages and whether the assessment data was sufficiently robust to allow progression.

Chapter 5. Discussion

In this instrumental case study, we explored the decision-making of two CC representing programs of different size and length. In both cases, the CC demonstrated a fixed mindset, with their meetings following a structured method for the majority of trainees and invoking a more extensive process of searching and sharing only when concerns were identified. The triggers for discussion consistently related to either CC members' uncertainty of the process to be followed or concerns with the adequacy of the data presented; the conversations that ensued contained attempts to clarify the process and understand the data in the context of the individual trainee and CC members' experiences. Our findings fell within the theoretical description of group process outlined by Hauer and colleagues (2021), and provide a case-based clarification of the relative importance of different processes for individual trainees. We present a theoretical framework illustrating the CC decision making process as we understood it, following our theory-informed inquiry (Figure 6.).

Competence Committee Decisions

The study of group process has led to many explanations of how small groups make decisions. The interplay of multiple theories led to the theoretical framework proposed by Hauer et al. (shown in Figure 4), which suggested that different aspects of committee function are explained by different theories (2021). Our study found that the CC followed a process that began within a social decision schema model and evolved to a discussion that invoked social influence theory, shared mental models, and social judgment scheme to clarify the points of contention. The decision making was at risk of

groupthink and premature closure, influenced by the group composition and the group and individual members' mindset.

Discussion and Information Sharing

The CC members in our study entered into discussions in order to improve their understanding of the situation. Studies of group decision making emphasize the importance of considering unshared information and the need to have a group leader who can identify members who are more likely to hold such information (Dennis, 1996; van Ginkel & van Knippenberg, 2009). The information presented to the CC could be considered to be the initially shared information; this was generally agreed upon as truth. When the presenter or another CC member identified a discrepancy in the data, the groups recognized that they required additional information to make a decision and began the search for unshared information. In both CC studied, when questions arose, the program directors were specifically invited to share information as they were recognized to have unique information and were trusted to have a sound understanding of the process.

Small groups make better decisions than individuals only when they share a conceptualization of the task at hand and the approach to be followed to reach their common goal (Edgar et al., 2021; Kerr & Tindale, 2004). This was exemplified in our study by the frequent questions about the requirements for making competence and progress decisions. The chair of CC1 began every meeting with a brief review of the terms of reference and functions of the committee. Despite this, questions about the relative importance of different types of data, number of EPAs required and timing of progression between stages of training were raised as frequently in CC1 as CC2. Their

task-specific shared mental model was therefore refined during each meeting and members reported feeling more confident in their role as a CC over time.

Data Adequacy

Challenges with the data were a major concern and common trigger of discussion in our study. These took the form of inadequate information, inconsistency between the numeric score and narrative comments, and misinterpretation of the complexity of EPA assessments by faculty assessors. The members interviewed were in agreement that their groups' decisions could only be correct if they had high quality data to work from. This was similar to a previous report that found discussions were frequently triggered by inadequate data (Pack et al., 2019). In that study, problematic data was typically described as incongruent or incomplete, and the committees most often chose to defer their decisions and search out more information.

Difficulty in obtaining useful assessment data and narrative comments was not a new problem in CBME, although the change from Likert scale scoring to entrustment scales was intended to help this process. Studies of the entrustment score used on current EPA assessments demonstrated that this approach was intuitive and reliable for raters (Gofton et al., 2012), however the interpretation of what each EPA intended to measure remained a challenge (Tomiak et al., 2020). Faculty development improved the information collected, and the comparison of multiple assessments from the same rater over time may be more meaningful to visualize progression, which contrasts with the general understanding that the validity of programmatic assessment is improved by the use of multiple assessors (Oudkerk Pool et al., 2018). Members interviewed in our study explained that in their consideration of problematic data, they would rely on the

narrative feedback to confirm the competence described in the numeric entrustment score as well as to confirm that the assessed task was interpreted as intended. It was apparent, though, that satisfactorily completing EPA assessments is neither sufficient nor necessary to be deemed competent. As trainees took different approaches to the achievement of EPAs, some are described as “gaming” the system (Pinsk et al., 2018), the number of EPA completed in a given time period may have reflected motivation or organization and planning, but was not thought to reflect competence. Complementing EPA data were in-training exams, 360 feedback and other sources of data; these were indispensable in trainee assessment. Members of both CC studied identified that trainee challenges were most often identified in 360 or holistic rotation feedback (for professionalism and communication), or on a knowledge exam rather than in the EPA data.

In agreement with previous publications (Ekpenyong et al., 2017; Lockyer et al., 2017; Oudkerk Pool et al., 2018), our study found that CC members formed a gestalt impression of trainee competence, based on multiple assessments of different types, by different raters, and considering their trajectory over a time period. This was in line with the intended role of CC as outlined by the RCPSC which emphasized the need to consider the whole of a trainee’s performance and not rely on the EPA data alone (*Competence committee guideline: process and procedures in decision making*, 2018). This reduced the impact of the problems with EPA data when it came to making competence decisions, but also raised a question as to the whether there was too much emphasis being placed on EPA assessments in the development of CBD training programs.

Mindset

The mindset of the CC members and the group as a whole affected their discussion of trainees. The Initial interpretation of trainee status was affected by the group mindset. The desire to discuss the trainees and identify target areas for improvement, reflecting a growth mindset or belief that every individual has the potential to improve (Dweck, 2016), was relevant to how frequently discussion triggers are raised. In an early description, CC were recognized to follow either a developmental or problem-identification model for resident review (Hauer et al., 2015); both CC in this study clearly followed the problem-identification model. The rapid identification of a trainee as struggling or performing well and choosing to discuss only the trainees in difficulty was a manifestation of a fixed mindset: that there is no need to try and help the high-achievers as they will continue to do well. The TOR for both CC studied included the goal of providing developmental guidance to every trainee. In the CC1 TOR this was delegated to the program director and assistant program director while the CC2 TOR did not specify. It was recognized that the CC review process can be lengthy and therefore committees have implemented methods to optimize the time spent on discussion (Chen et al., 2017). The CC members interviewed reported relying on the primary reviewer's take on the data due to time constraints and the volume of data that was in the assessment portfolios. In practice, both CC delegated the responsibility of providing developmental feedback to the PD and did not attempt to identify goals for residents who were progressing as expected, and only in a limited way for trainees in difficulty.

Committee members described the process of group decision-making as straightforward for high achieving residents. While they acknowledged that there was a risk to the rapid review of these residents, there was a feeling that CC meetings were not the forum for discussion on how to help them progress. Individual CC member mindset became more relevant in the discussion of individual trainees where it was evident that members wanted to identify the source of concern and help trainees improve. This was supported by the example of the discrepant high achiever presented above, where the discussion quickly ended once the problem was clarified as a process issue rather than a performance question. The conversation about specific trainees with problematic performance data included the assessment of their learning trajectory and a proposal of ideas for gaining additional learning experiences and mentorship for areas of concern. These ideas were still formulated into feedback for the learner by the PD, reflecting the integration of individual growth mindset within a problem-identification system.

Other Mediating Factors: Experience, Context and Data Interpretation

The CC members' individual perspectives and experience with trainees in the decision-making process were embedded in every discussion. In the meetings observed, this was apparent not as members sharing personal experiences and undocumented data, but rather their recognition that, "it's impossible to completely dissociate your own personal perspective having worked with them because personal memories and interactions are always much stronger than looking at numbers," (interview, F8). This was consistent with Kane's description that the interpretation of information takes place in the context of pre-existing assumptions (2001). According to

the social judgment scheme, members whose opinions are similar to one another will be considered of higher importance, and the information presented will be more likely to be accepted as true if it is consistent with the pre-formed opinion (Kerr & Tindale, 2004). This has also been described in CC function previously; assessors had a tendency to not change their impressions of trainees even when apparently conflicting evidence was presented (Oudkerk Pool et al., 2018). The impact of this source of bias was amplified in CC2 where all the members worked closely with the trainees and admitted to relying more on their impression of how the trainee was doing than the data presented.

Time pressures and a fixed mindset may also have led the groups toward premature closure, the tendency for reduced information sharing and early consensus seen when most members share the same pre-existing preference (Kerr & Tindale, 2004). Deconstructing the social judgment theory and groupthink identified strategies for more defensible decision making. The first proposes that the effect of novel information will be of greater impact when it is shared by one group member who is recognized as knowledgeable about the process (Kerr & Tindale, 2004). This was seen in our study as both CC1 and CC2 chairs attempted to redirect the conversation to invite sharing of additional data from the program directors when the discussion became subjective or off-track. The role of the program directors within the committee is interesting; as non-voting members they should have been there in advisory capacity as content and process experts, however, in CBD the CCs function as advisory to the PD in making final decisions about trainees and is responsible for putting in place remediation plans when needed. This identifies another source of bias with a risk of passing borderline candidates to avoid incurring additional work for the PD.

Next, the inclusion of members with diverse opinions and from varied contexts could increase sharing of novel information and perspectives (Surowiecki, 2005; Wittenbaum et al., 2004). The composition of the CC studied were both homogeneous, with a majority composition of clinical teaching faculty from within their specialty. While each CC included one member from an outside discipline, these individuals were not present at most meetings observed and represented a minority voice that may not have been sufficient to overcome the social decision structure. In establishing CC membership, programs should consider expanding the diversity of their committee with external members who have knowledge in the area of CBME or assessment, therefore they could be seen as process experts and respected even as a minority voice.

Feedback to Trainees

The conceptualization of CBME was of a learner-centred process designed to be individualized to each trainee's needs and rate of learning (Frank, Mungroo, et al., 2010). We found that there were two different challenges with respect to the resident contributions and impacts to this process. First, the CC in our study took into consideration the effects of their decisions on the trainees only when there were concerns identified. The majority of residents who were considered to be progressing as expected, and those who might have been progressing faster than expected did not receive developmental feedback from the committee or suggestions on how to further enrich their learning.

Second, CMBE required initiative from the trainees to seek observed encounters and assessment completion; this was undertaken to a variable extent by different trainees, which affected the completeness of their portfolios. This individual approach to

EPA collection was interpreted by some as a marker of competence while others believed it was more reflective of organizational skills and understanding of the CBME process (interviews, F2, F6, F12, F15). In either case, the trainees were expected to invest a lot of effort in the collection of assessment data which for many may not have led to meaningful feedback over time, given that there was little discussion about fostering further development for high-achieving residents. The CC that we studied did not feel that it would ever be reasonable to accelerate a trainee's progression, regardless of how exceptional or advanced they were in their performance.

The impact of CC meetings on residents has been described as anxiety-provoking, and that additional benefit to the residents could be gained by providing feedback on the discussion and recommendations for improvement (Rich et al., 2020). This was echoed in the findings of Pack (2020) who identified that the CC were in a unique position to provide robust data-driven feedback to trainees to guide further development. A systematic mechanism for return of recommendations that could be incorporated into personal learning plans would optimize the educational benefit of CBME and the CC process for trainees, in particular those who are high-achieving. As the majority of trainees were considered to be performing well and would not trigger discussion at a CC following a problem-identification model, it is possible that only a small fraction of trainees benefit from this redesign of the programme of assessment in residency training.

This study did not examine resident perspectives on the CC process or the impact they perceive CBD has on their learning. With the resident experience in mind and the constant need to balance the extraneous load placed on trainees, future study

should assess the feedback process following CC meetings and interrogate whether there is benefit to all trainees.

Limitations

This study was designed to delve into the functioning of two CC in a single institution, examining the process of 28 progress decisions. The findings are therefore reflective of the CC studied and may not be applicable to other programs or other institutions. Both programs were more medical specialties and their collection and use of workplace based assessment data is suspected to be different from surgical specialties. It is uncertain whether the extent of personal bias seen in our study would be present in a larger program where the CC members have less direct involvement with every trainee. Our goal was to provide insight into the process; the repetition of themes we identified in the two different programs, along with the parallels with other CC studies does increase the likelihood that our findings apply in other settings. While case study research is immersive and includes many sources of information, the trainee perspectives on CC were not included. This was due to the initial research question focus on decision making process; however, examining the downstream impact of those decisions would have added important insight for consideration.

Qualitative research is by its nature subjective and incorporates the perspectives of the researchers alongside the participants. Many of the CC members interviewed also had extensive experience within the field of medical education, thereby their responses may have been based on their knowledge of the intended function of CC within CBME rather than solely their experiences as CC members. This added to the

complexity and richness of the findings, as the semi-structured interviews included thoughtful reflections on CBME as a system as well as the functioning of their CC.

Future directions

Future work to build on these results may include investigation of the resident perspective on CC function and contribution to their learning, examination of the roles of mediating factors such as mindset, process uncertainty and data quality in other training programs, incorporation of our findings in faculty training tools for EPA assessments and competence committee feedback, among other ideas.

Chapter 6. Conclusion

The recommended role of a RCPSC post-graduate CC is to determine trainees' level of competence and readiness to progress through a training program, based on the assessment data available, and to provide feedback for learning to encourage their ongoing development. In this study, we found that CC made decisions about trainees based on the information provided, often with little discussion of the data or developmental recommendations. When problems were identified in the data or there was uncertainty around the process to be followed, CC members considered their own experiences and the trainee's context through a process of data exchange to reframe their shared understanding of the trainee's competence with respect to the decision to be made. This more elaborate process allowed for the group to be confident in the judgment they produced and provided more valuable feedback to the trainee, but is at risk of bias due to the implicit impressions of each CC member.

Our theory-informed analysis led to a theoretical framework of CC decision-making illustrating a process that began within a social decision schema model and evolved to a discussion that invokes social influence theory, shared mental models, and social judgment scheme to clarify the points of contention. This qualitative case study expanded on previously published theoretical models of decision making in CC and identified some areas for further consideration. We highlighted the role of personal experience and the limited discussion that occurs for trainees who are deemed to be doing well as considerations for refining the CC process to be more transparent and beneficial to stakeholders, including trainees and faculty members.

References

Andolsek, K., Padmore, J., Hauer, K. E., Edgar, L., & Holmboe, E. (2017). *Clinical competency committees: a guidebook for programs*. Accreditation Council for Graduate Medical Education Retrieved from https://www.acgme.org/Portals/0/ACGMEClinicalCompetencyCommitteeGuidebook.pdf

Bok, H. G., Teunissen, P. W., Favier, R. P., Rietbroek, N. J., Theyse, L. F., Brommer, H., . . . Jaarsma, D. A. (2013). Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Med Educ*, *13*, 123. https://doi.org/10.1186/1472-6920-13-123

Boulet, J. R., & Durning, S. J. (2018). What we measure ... and what we should measure in medical education. *Med Educ*. https://doi.org/10.1111/medu.13652

Chahine, S., Cristancho, S., Padgett, J., & Lingard, L. (2017). How do small groups make decisions?: a theoretical framework to inform the implementation and study of clinical competency committees. *Perspect Med Educ*, *6*(3), 192-198. https://doi.org/10.1007/s40037-017-0357-x

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE publications Ltd.

Charmaz, K. (2008). Constructionism and the grounded theory. In J. A. Holstein & J. F. Gubrium (Eds.), *Handbook of constructionist research* (pp. 397-412). The Guilford Press.

Charters, E. (2003). The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods. *Brock education*, *12*(2). https://doi.org/10.26522/brocked.v12i2.38

Chen, F., Arora, H., & Martinelli, S. M. (2017). Use of key performance indicators to improve milestone assessment in semi-annual clinical competency committee meetings. *The Journal of Education in Perioperative Medicine JEPM*, *19*(4), E611.

Competence by Design cheat sheet. (2016). The Royal College of Physicians and Surgeons of Canada Retrieved from http://www.royalcollege.ca/rcsite/documents/cbd/cbd-cheat-sheet-e.pdf

Competence committee guideline: process and procedures in decision making. (2018). The Royal College of Physicians and Surgeons of Canada Retrieved from https://www.royalcollege.ca/rcsite/documents/cbd/competence-committees-process-procedures-e

Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC Med Res Methodol*, *11*, 100. https://doi.org/10.1186/1471-2288-11-100

Dennis, A. R. (1996). Information Exchange and Use in Small Group Decision Making. *Small group research*, *27*(4), 532-550. https://doi.org/10.1177/1046496496274003

Donato, A. A., Alweis, R., & Wenderoth, S. (2016). Design of a clinical competency committee to maximize formative feedback. *Journal of Community Hospital*

Internal Medicine Perspectives, *6*(6), 33533. https://doi.org/10.3402/jchimp.v6.33533

Duijn, C. C. M. A., Welink, L. S., Bok, H. G. J., & Ten Cate, O. T. J. (2018). When to trust our learners? Clinical teachers' perceptions of decision variables in the entrustment process. *Perspect Med Educ*, *7*(3), 192-199. https://doi.org/10.1007/s40037-018-0430-0

Dweck, C. (2016). *Mindset: The new psychology of success*. Ballantine Books.

Edgar, L., Jones, M. D., Jr., Harsy, B., Passiment, M., & Hauer, K. E. (2021). Better Decision-Making: Shared Mental Models and the Clinical Competency Committee. *J Grad Med Educ*, *13*(2 Suppl), 51-58. https://doi.org/10.4300/jgme-d-20-00850.1

Ekpenyong, A., Baker, E., Harris, I., Tekian, A., Abrams, R., Reddy, S., & Park, Y. S. (2017). 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*, *39*(10), 1074-1083. https://doi.org/10.1080/0142159x.2017.1353070

Frank, A. K., O'Sullivan, P., Mills, L. M., Muller-Juge, V., & Hauer, K. E. (2019). Clerkship Grading Committees: the Impact of Group Decision-Making for Clerkship Grading. *J Gen Intern Med*, *34*(5), 669-676. https://doi.org/10.1007/s11606-019-04879-x

Frank, J. R., Mungroo, R., Ahmad, Y., Wang, M., De Rossi, S., & Horsley, T. (2010). Toward a definition of competency-based education in medicine: a systematic

review of published definitions. *Med Teach*, *32*(8), 631-637. https://doi.org/10.3109/0142159X.2010.500898

Frank, J. R., Snell, L., & Sherbino, J. (Eds.). (2015). *CanMEDS 2015 physician competence framework*. Royal College of Physicians and Surgeons of Canada.

Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R., . . . Harris, K. A. (2010). Competency-based medical education: theory to practice. *Medical Teacher*, *32*(8), 638-645. https://doi.org/10.3109/0142159X.2010.501190

Gaglione, M. M., Moores, L., Pangaro, L., & Hemmer, P. A. (2005). Does group discussion of student clerkship performance at an education committee affect an individual committee member's decisions? *Academic Medicine*, *80(10)*(Supplement), S55-S58.

Gofton, W., Dudek, N., Barton, G., & Bhanji, F. (2017). *Assessment implementation guide: formative tips for medical teaching practice*. Ottawa: The Royal Colege of Physicians and Surgeons of Canada Retrieved from http://www.royalcollege.ca/rcsite/documents/cbd/wba-implementation-guide-tips-medical-teaching-practice-e.pdf

Gofton, W. T., Dudek, N. L., Wood, T. J., Balaa, F., & Hamstra, S. J. (2012). The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. *Academic Medicine*, *87*(10), 1401-1407. https://doi.org/10.1097/ACM.0b013e3182677805

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough?:An Experiment with Data Saturation and Variability. *Field Methods*, *18*(1), 59-82. https://doi.org/10.1177/1525822x05279903

Hall, A. K., Rich, J., Dagnone, J., Weersink, K., Caudle, J., Sherbino, J., . . . Van Melle, E. (2018). P061: Implementing CBME in emergency medicine: lessons learned from the first 6 months of transition at Queens University. *Canadian Journal of Emergency Medicine*, *20*, S78. https://doi.org/10.1017/cem.2018.259

Hatala, R., Sawatsky, A. P., Dudek, N., Ginsburg, S., & Cook, D. A. (2017). Using In-Training Evaluation Report (ITER) Qualitative Comments to Assess Medical Students and Residents: A Systematic Review. *Acad Med*, *92*(6), 868-879. https://doi.org/10.1097/ACM.0000000000001506

Hauer, K. E., Cate, O. T., Boscardin, C. K., Iobst, W., Holmboe, E. S., Chesluk, B., . . . O'Sullivan, P. S. (2016). Ensuring resident competence: a narrative review of the literature on group decision making to inform the work of clinical competency committees. *J Grad Med Educ*, *8*(2), 156-164. https://doi.org/10.4300/jgme-d-15-00144.1

Hauer, K. E., Chesluk, B., Iobst, W., Holmboe, E., Baron, R. B., Boscardin, C. K., . . . O'Sullivan, P. S. (2015). Reviewing residents' competence: a qualitative study of the role of clinical competency committees in performance assessment. *Acad Med*, *90*(8), 1084-1092. https://doi.org/10.1097/acm.0000000000000736

Hauer, K. E., Edgar, L., Hogan, S. O., Kinnear, B., & Warm, E. (2021). The Science of Effective Group Process: Lessons for Clinical Competency Committees. *J Grad Med Educ*, *13*(2 Suppl), 59-64. https://doi.org/10.4300/jgme-d-20-00827.1

Hemmer, P. A., & Kelly, W. F. (2017). We need to talk: clinical competency committees in the key of c(onversation). *Perspect Med Educ*, *6*(3), 141-143. https://doi.org/10.1007/s40037-017-0360-2

Holmboe, E. S., Sherbino, J., Long, D. M., Swing, S. R., & Frank, J. R. (2010). The role of assessment in competency-based medical education [Article]. *Medical Teacher*, *32*(8), 676-682. https://doi.org/10.3109/0142159X.2010.500704

Iobst, W. F., Sherbino, J., Cate, O. T., Richardson, D. L., Dath, D., Swing, S. R., . . . Frank, J. R. (2010). Competency-based medical education in postgraduate medical education [Article]. *Medical Teacher*, *32*(8), 651-656. https://doi.org/10.3109/0142159X.2010.500709

Kane, M. T. (2001). Current Concerns in Validity Theory. *Journal of educational measurement*, *38*(4), 319-342. https://doi.org/10.1111/j.1745-3984.2001.tb01130.x

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annu Rev Psychol*, *55*(1), 623-655. https://doi.org/10.1146/annurev.psych.55.090902.142009

Klutts, J. S., Guerin, L. A., Bruch, L. A., Firchau, D. J., Knudson, C. M., Rosenthal, N. S., . . . Krasowski, M. D. (2015). Pathology milestones: assessing clinical competency by committee. *Academic Pathology*, *2*(4), 2374289515614003. https://doi.org/10.1177/2374289515614003

Lockyer, J., Carraccio, C., Chan, M. K., Hart, D., Smee, S., Touchie, C., . . . Frank, J. R. (2017). Core principles of assessment in competency-based medical education. *Med Teach*, *39*(6), 609-616. https://doi.org/10.1080/0142159x.2017.1315082

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Acad Med*, *65*(9 Suppl), S63-67.

Moser, A., & Korstjens, I. (2018). Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis. *The European journal of general practice*, *24*(1), 9-18. https://doi.org/10.1080/13814788.2017.1375091

Nasca, T. J., Philibert, I., Brigham, T., & Flynn, T. C. (2012). The Next GME Accreditation System — Rationale and Benefits. *N Engl J Med*, *366*(11), 1051-1056. https://doi.org/10.1056/NEJMsr1200117

Oudkerk Pool, A., Govaerts, M. J. B., Jaarsma, D., & Driessen, E. W. (2018). From aggregation to interpretation: how assessors judge complex data in a competency-based portfolio. *Adv Health Sci Educ Theory Pract*, *23*(2), 275-287. https://doi.org/10.1007/s10459-017-9793-y

Pack, R., Lingard, L., Watling, C., & Cristancho, S. (2020). Beyond summative decision making: Illuminating the broader roles of competence committees. *Med Educ*, *54*(6), 517-527. https://doi.org/10.1111/medu.14072

Pack, R., Lingard, L., Watling, C. J., Chahine, S., & Cristancho, S. M. (2019). Some assembly required: tracing the interpretative work of Clinical Competency Committees. *Med Educ*. https://doi.org/10.1111/medu.13884

Pinsk, M., Karpinski, J., & Carlisle, E. (2018). Introduction of Competence by Design to Canadian Nephrology Postgraduate Training. *Canadian journal of kidney health and disease*, *5*, 2054358118786972-2054358118786972. https://doi.org/10.1177/2054358118786972

Rekman, J., Gofton, W., Dudek, N., Gofton, T., & Hamstra, S. J. (2016). Entrustability Scales: Outlining Their Usefulness for Competency-Based Clinical Assessment. *Acad Med*, *91*(2), 186-190. https://doi.org/10.1097/ACM.0000000000001045

Rich, J. V., Fostaty Young, S., Donnelly, C., Hall, A. K., Dagnone, J. D., Weersink, K., . . . Klinger, D. A. (2020). Competency-based education calls for programmatic assessment: But what does this look like in practice? *J Eval Clin Pract*, *26*(4), 1087-1095. https://doi.org/10.1111/jep.13328

Schuwirth, L., & Ash, J. (2013). Assessing tomorrow's learners: in competency-based education only a radically different holistic method of assessment will work. Six things we could forget [Article]. *Medical Teacher*, *35*(7), 555-559. https://doi.org/10.3109/0142159X.2013.787140

Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor Books.

Swing, S. R. (2007). The ACGME outcome project: retrospective and prospective. *Med Teach*, *29*(7), 648-654. https://doi.org/10.1080/01421590701392903

Tekian, A., Hodges, B. D., Roberts, T. E., Schuwirth, L., & Norcini, J. (2015). Assessing competencies using milestones along the way. *Med Teach*, *37*(4), 399-402. https://doi.org/10.3109/0142159x.2014.993954

ten Cate, O. (2005). Entrustability of professional activities and competency-based training. *Med Educ*, *39*(12), 1176-1177. https://doi.org/10.1111/j.1365-2929.2005.02341.x

ten Cate, O., Hart, D., Ankel, F., Busari, J., Englander, R., Glasgow, N., . . . Wycliffe-Jones, K. (2016). Entrustment decision making in clinical training. *Acad Med*, *91*(2), 191-198. https://doi.org/10.1097/acm.0000000000001044

Thornberg, R., & Charmaz, K. (2013). Grounded theory and theoretical coding. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 153-169). SAGE Publications Ltd. https://doi.org/10.4135/9781446282243

Tomiak, A., Braund, H., Egan, R., Dalgarno, N., Emack, J., Reid, M. A., & Hammad, N. (2020). Exploring How the New Entrustable Professional Activity Assessment Tools Affect the Quality of Feedback Given to Medical Oncology Residents. *J Cancer Educ*, *35*(1), 165-177. https://doi.org/10.1007/s13187-018-1456-z

van der Vleuten, C. P., Schuwirth, L. W., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K., & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Med Teach*, *34*(3), 205-214. https://doi.org/10.3109/0142159x.2012.652239

van Ginkel, W. P., & van Knippenberg, D. (2009). Knowledge about the distribution of information and group decision making: When and why does it work? *Organizational behavior and human decision processes*, *108*(2), 218-229. https://doi.org/10.1016/j.obhdp.2008.10.003

Williams, R. G., Schwind, C. J., Dunnington, G. L., Fortune, J., Rogers, D., & Boehler, M. (2005). The effects of group dynamics on resident progress committee deliberations. *Teach Learn Med*, *17*(2), 96-100. https://doi.org/10.1207/s15328015tlm1702_1

Wittenbaum, G. M., Hollingshead, A. B., Paulus, P. B., Hirokawa, R. Y., Ancona, D. G., Peterson, R. S., . . . Yoon, K. (2004). The Functional Perspective as a Lens for Understanding Groups. *Small group research*, *35*(1), 17-43. https://doi.org/10.1177/1046496403259459

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6 ed.). SAGE Publications, Inc.

Appendix A

Interview guide

Introduction

Thank you for participating in this interview. We appreciate you taking the time to share your views and impressions of the competence committee functioning. We are interested in the methods by which competence committees and their members make decisions about residents. This interview is being recorded so that it can be transcribed verbatim for qualitative analysis. Please do not identify residents or committee members by name during the interview; if this occurs, the transcriptionist will be instructed to remove and replace with "NAME".

It is possible that during review of the interview data, clarification might be needed and you could be contacted for a brief follow-up to verify that our interpretation aligns with what you intended to say.

Interview questions

1. How do you as a committee member make decisions about residents?
    Additional prompts:
        Prior to the meeting, which materials were most helpful in formulating your impression of the resident?
        What is the most challenging aspect of making decisions about whether the resident is at the appropriate level using this process?
        How do you weigh your own experience with the learner when you are working with the committee and assessment data?

2. Please tell me about how the competence committee makes decisions about residents.
    Additional prompts:
        What is done prior to the meeting, by whom, and what happens at the meeting?
        How is the process different for learners who are doing well vs. struggling?
        How is the process different for learners at transition points?
        How does the committee come to an agreement if there are strong differences of opinion?
        What would make the competence committee process easier?
        How has the use of virtual meeting technology affected your process?[1]

3. How does the committee use the resident assessments to inform their decisions?
    Additional prompts:
        What happens if the numeric score doesn't match the comments?

[1] Question added due to the COVID-19 pandemic changing meeting structure

How do you interpret narrative comments? Ambiguous or non-specific comments? Divergent comments?

4. Describe the group dynamics of your committee
   Additional prompts:
   Tell me about the power dynamics among the group?
   Is there a dominant voice at the committee? IF yes – Why is that voice dominant? How does that voice influence other members' decision making?
   What is the influence of the opinions of others and the discussion at the committee on your initial impression of the resident being reviewed?
   When members of the committee disagree about the resident, what are the reasons?
   How does the varied experiences of the committee members with residents in general and with the individuals being assessed contribute to the discussion?

5. How do you feel about the ability of the CC to assess resident ability and progression?
   Additional prompts:
   How would you describe the main role of your committee? How does this role differ between residents who excel and those who struggle?
   How does the CC identify residents in difficulty?

6. How does the CC consider the context requirements of the EPAs?[2]

7. How does the CC address the number of in progress assessments as compared to a resident who has only successful EPA observations?
   Does the CC show a growth mindset?

8. What would it take to hold someone back?[3]
   Would you consider rating a trainee as "progressing faster than expected"?

9. How does the CC decision get communicated back to the trainee?
   Is it different for those doing well vs those who are struggling?

10. Does the EPA count and progress through the CBD structure reflect true differences between trainees?[4]

[2] Questions 6 and 7 were added after 2nd committee meeting, for 2nd round of interviews
[3] Questions 8 and 9 were added after 3rd committee meeting, for 3rd round of interviews
[4] Question 10 was added after 3rd round of interviews

Appendix B

CC1 Competence Committee Terms of Reference (Redacted)

Purpose: The purpose of the CC1 Competence committee (CC) is to regularly review resident files with the goal of ensuring that learners achieve all the requirements of the discipline as defined by the Royal College of Physicians and Surgeons of Canada. The CC makes recommendations regarding learner status within the Royal College Residency Program at the University; however, any decisions regarding remediation or intent to withdraw require ratification by the Residency Training Committee (RTC).

The goal of the CC is to ensure that all learners achieve the requirements of the discipline. The CC makes decisions and recommendations by integrating data from the various tools used in our program of assessment including entrustable professional activity (EPA) observations, daily assessments, off service rotation assessments, quarterly and national in-training exams as well the other assessment tools. The CC process allows for an informed group decision-making process where patterns of performance can be collated to reveal a broad picture of a resident's progression toward competence.

A regular review of residents' progress facilitates a developmental approach, supporting resident learning over time. The CC should help the resident education team identify residents who are not meeting their milestones and can suggest or mandate support and coaching for the resident before the resident gets too far off their trajectory. This support can mean, for example, assigning special mentors, extra readings or modified rotations.

In Competence-by-Design (CBD), promotion recommendations are made by the competence committee away from the individual teacher-learner interactions. By shifting broader promotion discussions to the CC, interactions between individual residents and observers can focus on coaching feedback to help improve residents' performance (i.e. assessment for learning).

The CC's purpose is to determine if residents have met the appropriate standard, or are on an appropriate trajectory, to move between stages on the competence continuum and to determine when residents are ready for the Royal College examinations, as well as Certification upon completion of their Transition to Practice (TTP) stage of training.

The CC is expected to exercise judgement in making progress decisions: i.e. they use the Royal College Specialty Committee suggested number of EPA observations as a guideline but are not bound to a specific number of assessments. In addition to using EPAs, the committee decisions will be based on all assessment tools and relevant evidence uploaded into the resident file. Individual committee member experience can only be introduced with appropriate documentation in the resident file.

The committee must feel it has adequate information to make holistic judgements on the progression of the resident.

All committee discussions are strictly confidential and only shared on a professional need-to-know basis. This principle is equivalent to patient confidentiality in clinical medicine.

All committee decisions are to be made in a spirit of supporting each trainee in achieving their own individual progression of competence. The CC has a responsibility to make decisions in the spirit of protecting patients from harm, including weighing a trainees' progress in terms of what they can safely be entrusted to perform with indirect supervision. Some CCC discussions must be shared to provide focused support and guidance for the trainee.

The CC, on an exceptional basis, may have the option to identify trainees who are eligible for an accelerated learning pathway.

Likewise, on an exceptional basis and after due process, the CC has the responsibility to identify the trainees who have met the predefined category of *failure to progress*, and who should be requested to leave the program (see the University Postgraduate Medical Education (PGME) remediation, probation and dismissal policy).

Decisions regarding accelerated learning pathway or failure to progress must be presented to the RTC by the program director (PD) and approved by the RTC.

Members:

The CC is made up of 6 voting staff physicians, one of whom is from a department outside of the specialty and 3 non-voting (communicating) members (the program director, assistant program director, and resident advocate).

The CC is Chaired by a member of the clinical teaching faculty and someone other than the program director. As a member (not Chair) of the competence committee, the program director can fully participate in the discussions and not be hampered by the need to wear multiple hats during the discussion.

<u>Chair:</u>

The CC has 6 voting members one of whom will act as chair and 5 of whom will act as reviewers. The program director or assistant program director cannot be the Chair. The Chair will set the agenda, in conjunction with the program director or their delegate, and maintain order of the meetings. The Chair will welcome members, present the agenda and orient the members to the decisions that will be made at the meeting. The Chair will remind members of the confidentially policy prior to the start of the meeting. The Chair will ensure the meeting is run efficiently and that all trainees on the agenda are discussed. The Chair will direct discussion, keeping reviewers on task, and moving toward consensus. When

consensus of the members cannot be reached, the Chair can move to postpone the decisions (see guideline below on process for the meeting). On completion of the meeting the Chair will ensure that documentation of the status of the trainee is shared with the RTC (if required), program director, and resident file.

Primary Reviewer:

Each trainee selected for discussion at the CC meeting is assigned a designated primary reviewer. The primary reviewer is responsible to complete a detailed review of progress of the trainee based on all assessment tools, observations, completed EPAs, and preceptor feedback documented in the resident file. The primary reviewer will provide a succinct synthesis and impression of the trainee's progress to the CC. The primary reviewer will then propose a resolution on that trainee's status going forward. If there is concern raised about resident performance by the primary reviewer or if the reviewer is unsure what status to propose to the committee, they may request a second review by another member of the committee. They will then jointly present the trainee at the CC meeting.

Resident Advocate:

The resident advocate will attend the CC in the role of a nonvoting member. Their primary responsibility will be to ensure the viewpoint needs of the resident are respected and cared for.

Term of Office:

The CC members will be invited to participate at the discretion of the program director. CC members are encouraged to participate in a 3-5 year term. Renewal of a CC member will occur yearly as determined by the program director. The Chair will be determined yearly and may be delegated to another committee member for an individual meeting.

Meetings:

The CC meets at least twice per year; however, the goal is to meet every quarter (mid to end September, mid-December, March and middle of June). Every trainee in the program must be discussed a minimum of twice per year.

During the meeting each trainee is considered in turn, with the primary reviewer presenting their synthesis, displaying relevant reports and sharing important quotes from comments recorded in the trainee's file. The primary reviewer concludes by proposing a *status* for the trainee going forward in the program and an action. The status can be any of those in Table 3.

If the proposed status of the trainee is seconded by another committee member, all members are invited to discuss the motion. The Chair will call a vote to the proposed recommendation of the primary reviewer. If the recommendation is not seconded or the motion does not achieve a majority of votes, the Chair will then request another motion regarding the trainee. This process will continue until a majority of the committee

members support the status motion. If a member is a longitudinal preceptor or resident coach for a trainee, they will not participate in the status vote

A status decision can only be deferred if additional information is required. The deferred decision must be revisited within 4 weeks.

Table 3. *Possible Learner Status and Actions for CC*

Learner Status	Learner - Resident Action
Progressing as Expected	Monitor Learner
	Modify Learning Plan – Suggested Focus on EPA/IM observations or RTE
	Promote Learner to Stage 2
	Promote Learner to Stage 3
	Promote Learner RC Exam eligible
	Promote Learner to Stage 4
	Promote Learner RC Certification Eligible
Not Progressing As Expected	Modify Learning Plan – Suggested Focus on EPA/IM observations or RTE
	Formal Remediation
Progress Is Accelerated	Modify Learning Plan – Suggested Focus on EPA/IM observations or RTE
	Promote Learner to Stage 2
	Promote Learner to Stage 3
	Promote Learner RC Exam eligible
	Promote Learner to Stage 4
	Promote Learner RC Certification Eligible
Failure to Progress	Modify Learning Plan – Additional Focus on EPA/IM observations or RTE
	Formal Remediation
	Withdraw Training
Inactive	Monitor Learning – Resident (expected return – parental leave, sick leave etc)
	Withdraw training

Note: IM individual milestone, RTE required training experience

Agenda development:

Trainees are selected for a planned CC meeting by the Chair of the committee, the program director, or their delegate. This scheduling occurs in advance of the committee meeting so that assigned reviewers (see below) have sufficient time to review the resident file prior to the meeting.

Trainee Selection for the meeting:

Trainees may be selected for review based on:

- A regularly timed review.
- A concern has been flagged.
- Completion of stage requirements and eligible for promotion or completion of training.
- Readiness for the Royal College exam.
- A significant delay or acceleration in the trainee's progress or academic performance.

Quorum:

There should be at least 50% of the competence committee in attendance with a minimum of 3 voting members of the competence committee to achieve quorum. Either the program director or assistant program director must be present for all discussions.

Reporting:

The outcomes of the CC discussions are recorded in the minutes of the CC. The status decision for the trainee must be communicated by the program director to the trainee in a timely manner. CC decisions regarding learner status must be documented in the resident file.

Changes to the trainee’s learning plan, assessments, or rotation schedule will be developed with the trainee, the program director, the assistant program director, and the trainee’s longitudinal preceptor. These adjustments must be implemented as soon as feasible.

Appeals Process:

Resident appeals should comply with the University PGME Resident Appeals Policy. An appeal of the CC decision can be made to the RTC.

Appendix C

CC2 Competence Committee Terms of Reference (Redacted)

Purpose:

The purpose of the CC2 Competence Committee is to review resident performance in a robust and transparent manner and make decisions regarding the progress of residents in the Program in achieving the national standards established by the Royal College of Physicians and Surgeons of Canada (RCPSC) Specialty Committee.

Role:

The Competence Committee allows for an informed group decision-making process where patterns of performance can be collated to reveal a broad picture of a resident's progression toward competence.

The mandate of the Competence Committee is to review and discuss learner portfolios in order to:

- Advise and guide resident learning and development;
- Adjust a resident's training experiences to enhance learning opportunities;
- Review assessments to determine a resident's achievement of each Entrustable Professional Activity (EPA);
- Recommend learner status changes and progression of residents through stages of training based on achievement of EPAs.
- Ensure there is a mechanism by which residents are informed of status changes following a review.

Responsibility and Authority:

The Competence Committee reports to the RPC via the Program Director and will be responsible for monitoring, guiding and recommending necessary training experiences to help residents in the residency program achieve required competencies. This is achieved by

- Monitoring and making decisions on the progress of each resident in demonstrating achievement of the EPAs or independent milestones within each stage of the competency based residency program.
- Guiding resident training experiences and learning plans to facilitate achievement of EPAs or independent milestones.
- Synthesizing the results from multiple assessments and observations (including EPAs, simulation activities, scholarly project reports, and other training experiences or approved formal evaluations) to make recommendations to the RPC related to:
    - The promotion of residents to the next stage of training;
    - The review and approval of individual learning plans developed to address areas for improvement. This may include rotation modifications, formal course attendance and specific mentorship arrangements;
    - Determining readiness to challenge the Royal College examinations;

    - Determining readiness to enter independent practice on completion of the transition to practice stage;
    - Determining that a resident is failing to progress within the program;
    - Monitoring the outcome of any learning or improvement plan established for an individual resident;
- Maintaining confidentiality and promoting trust by sharing information only with individuals directly involved in the development or implementation of learning or improvement plans.

Composition:

The Competence Committee will be composed of clinical teaching faculty affiliated with the Residency Training Program. The Competence Committee will be chaired by one of the voting members, who is not the program director.

An Academic Advisor will be assigned to each resident and will report resident progress to the Competence Committee. The Academic Advisor will sit on the Competence Committee, but will be a non-voting member for the resident he/she advises.

Non-voting members

- Program director
- Academic Advisor (cannot vote on resident he/she advises)
- Program administrator

Voting members

- Competence Committee chair
- Clinical teaching faculty of the RPC

Key Competencies and Characteristics:

The Competence Committee will be composed of individuals with interest, experience and expertise in assessment and medical education relevant to the specialty.

The Competence Committee members must be able to interpret multiple sources of qualitative and quantitative observation data to achieve consensus, where possible, in order to make judgments on outcomes.

Reporting:

The Competence Committee will report outcomes of discussions and decisions to the RPC.

Term of Office:

The selection of members of the Competence Committee will be based on established university policies. Members will be appointed by the Program Director to serve for 3 years, with possibility of renewal.

Meetings and Quorum:

- The Competence Committee will meet at minimum 4 times per year.
- Each resident will be discussed a minimum of 2 times per year. More frequent meetings may be required in the event that a resident is struggling to achieve EPAs or independent milestones. Extra meetings may be called on an ad hoc basis by the chair.
- Competence Committee members may attend meetings in person and/or via tele/video-conferencing.
- Program director or delegate is present at all meetings
- Quorum will be established when a minimum of 3 clinical preceptors are available and 50% of the committee is present.

Process

1. Agenda

Residents will be selected for review based on:

- Minimum frequency timing
- Performance concerns
- Eligibility for promotion
- Significant delay or acceleration in a trainee's progress

The resident will be informed of the specific purpose of the review and asked to collate all documented evaluations

2. Primary Reviewer

The resident's Academic Advisor or delegate, will be the primary reviewer

The resident and his/her Academic Advisor will work together to summarize all evaluations and make a proposal to the committee based on the specific purpose of the review. Should any deficits be identified by the resident and the Academic Advisor, the proposal should include a plan for rectifying those concerns.

At the Competence Committee meeting, the resident's Academic Advisor will present the proposal submitted by the resident, review the resident's overall progress and make a status decision recommendation to the Competence Committee.

Status decisions may include any of the following:

- Full promotion to next stage of training
- Progressing as expected within a stage of training
- Ready to challenge the Royal college examination
- Ready to enter independent practice on completion of transition to practice (TTP)
- Not progression as expected with minor concerns noted – development of an individual learning plan
- Failure to Progress with major concern noted – formal remediation recommended

- Failure to progress to next stage of residency training – this can only be advised after an unsuccessful remediation process
- Advise withdrawal from the residency training program – this can only be advised after a formal unsuccessful probation process

3. Secondary Reviewers

All other Competence Committee members are responsible for reviewing all residents on the agenda as secondary reviewers. All secondary reviewers are required to come prepared to discuss the progress of all trainees on the agenda.

4. Procedure

- The Chair welcomes all members and lists the residents who are intended for review, reminding members of the confidentiality of the proceedings.
- Each resident is discussed in turn with the Academic Advisor presenting the resident.
- The Chair will call a vote on the proposed status decision of the primary reviewer. If the motion does not achieve majority, the Chair will then request another motion. This process will continue until the majority of Competence Committee members support a status decision.
- A status decision can only be deferred if additional information is required. The deferred decision must be revisited within 4 weeks.
- All recommendations made by the Competence Committee will be documented in the resident's ePortfolio, recorded in the committee's meeting minutes, and presented to the RPC for ratification.
- The Competence Committee will endeavor to flag EPAs or Milestones which are inconsistently met at a defined stage for a cohort of residents to the Program Director. The Program Director, in turn, and in conjunction with the RPC, should alert the Specialty Committee to determine the appropriateness and expected time of completion of those EPAs.

5. Post Competence Committee Meeting

- Once the status decision is ratified by the RPC, the Program Director will communicate the decision to the resident within the next 1-2 weeks, ideally via an in-person or video conference meeting but alternatively by phone or email if the committee ratifies the Academic Advisor's proposed status decision.
- If applicable, the Program Director will use this meeting to discuss changes to the trainee's learning plan, assessments, or rotation schedule.

6. Appeal Process

- Competence Committee recommendations may be appealed in accordance with University Postgraduate Medical Education (PGME) Appeals policies.