How Do Clinical Competency Committees Assess Resident Performance on the Internal Medicine Milestones?

BY

ANDEM EKPENYONG
B.S., Loyola University, 1992
M.D., Rush Medical College, 1996

THESIS
Submitted as partial fulfillment of the requirements for the degree of Master in Health Professions Education in the Graduate College of the University of Illinois at Chicago, 2016

Chicago, Illinois

Defense Committee:

Yoon Soo Park, Chair and Advisor
Ilene Harris
Ara Tekian
Richard Abrams, Rush University Medical Center
ACKNOWLEDGMENTS

I would like to thank my thesis committee (Yoon Soo Park, Ilene Harris, Ara Tekian and Richard Abrams) for their guidance and mentorship. I would also like to thank the RUMC internal medicine residency program coordinators (Emelina Quintana and Laney McDougal) for their support of this project. In addition, I am grateful to our study participants, the RUMC internal medicine residents from whom I continue to learn, and the RUMC internal medicine department leadership for their support via the academic career development program.

AE
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION .................................................................</td>
<td>1</td>
</tr>
<tr>
<td>A. Background .................................................................</td>
<td>1</td>
</tr>
<tr>
<td>B. Statement of the Problem ...............................................</td>
<td>1</td>
</tr>
<tr>
<td>C. Purpose of the Study ...................................................</td>
<td>2</td>
</tr>
<tr>
<td>II. METHODS ........................................................................</td>
<td>4</td>
</tr>
<tr>
<td>A. Setting and Participants ...............................................</td>
<td>4</td>
</tr>
<tr>
<td>B. RUMC CCC Structure and Process .....................................</td>
<td>5</td>
</tr>
<tr>
<td>C. Study Design/Data Analysis ..........................................</td>
<td>6</td>
</tr>
<tr>
<td>1. Quantitative data .........................................................</td>
<td>6</td>
</tr>
<tr>
<td>2. Qualitative data ...........................................................</td>
<td>6</td>
</tr>
<tr>
<td>III. RESULTS .......................................................................</td>
<td>10</td>
</tr>
<tr>
<td>A. Quantitative .................................................................</td>
<td>10</td>
</tr>
<tr>
<td>B. Qualitative .................................................................</td>
<td>10</td>
</tr>
<tr>
<td>1. Data input .................................................................</td>
<td>11</td>
</tr>
<tr>
<td>2. CCC processing issues ................................................</td>
<td>12</td>
</tr>
<tr>
<td>3. Data output .................................................................</td>
<td>15</td>
</tr>
<tr>
<td>IV. DISCUSSION ....................................................................</td>
<td>19</td>
</tr>
<tr>
<td>A. Context .....................................................................</td>
<td>20</td>
</tr>
<tr>
<td>B. Strengths and Limitations ............................................</td>
<td>22</td>
</tr>
<tr>
<td>C. Implications .................................................................</td>
<td>23</td>
</tr>
<tr>
<td>V. CONCLUSION ..................................................................</td>
<td>25</td>
</tr>
<tr>
<td>APPENDICES .....................................................................</td>
<td>33</td>
</tr>
<tr>
<td>Appendix A ..................................................................</td>
<td>33</td>
</tr>
<tr>
<td>Appendix B ..................................................................</td>
<td>36</td>
</tr>
<tr>
<td>CITED LITERATURE ...............................................................</td>
<td>39</td>
</tr>
<tr>
<td>BIBLIOGRAPHY ..................................................................</td>
<td>41</td>
</tr>
<tr>
<td>VITA ..........................................................................</td>
<td>43</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. DESCRIPTIVE DATA FOR 2 REPORTING CYCLES (DECEMBER 2014 AND JUNE 2015)</td>
<td>26</td>
</tr>
<tr>
<td>II. FREQUENCY OF “NOT ABLE TO ASSESS (NA) RATINGS OVER 2 REPORTING CYCLES- DEC 2014 AND JUN 2015</td>
<td>27</td>
</tr>
<tr>
<td>III. MAJOR THEMES AND SUBTHEMES FROM CCC MEMBERS</td>
<td>28</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mean relative weights (%) of different types of assessments: Data aggregated over reporting periods (Mean ± 95% Confidence Intervals)</td>
<td>30</td>
</tr>
<tr>
<td>2. Milestones levels by core competency for each postgraduate year levels across two reporting periods: Mean ± 95% confidence intervals</td>
<td>31</td>
</tr>
<tr>
<td>3. Factors affecting data flow through the CCC</td>
<td>32</td>
</tr>
</tbody>
</table>
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACGME</td>
<td>Accreditation Council of Graduate Medical Education</td>
</tr>
<tr>
<td>CCC</td>
<td>Clinical Competency Committee</td>
</tr>
<tr>
<td>ICS1</td>
<td>Communicates effectively with patients and caregivers</td>
</tr>
<tr>
<td>ICS2</td>
<td>Communicates effectively in interprofessional teams</td>
</tr>
<tr>
<td>ICS3</td>
<td>Appropriate utilization and completion of health records</td>
</tr>
<tr>
<td>MK1</td>
<td>Clinical knowledge</td>
</tr>
<tr>
<td>MK2</td>
<td>Knowledge of diagnostic testing and procedures</td>
</tr>
<tr>
<td>NA</td>
<td>Not Able to Assess</td>
</tr>
<tr>
<td>NAS</td>
<td>Next Accreditation System</td>
</tr>
<tr>
<td>PBLI1</td>
<td>Monitors practice with a goal for improvement</td>
</tr>
<tr>
<td>PBLI2</td>
<td>Learns and improves via performance audit</td>
</tr>
<tr>
<td>PBLI3</td>
<td>Learns and improves via feedback</td>
</tr>
<tr>
<td>PBLI4</td>
<td>Learns and improves at the point of care</td>
</tr>
<tr>
<td>PC1</td>
<td>Gathers and synthesizes essential and accurate information to define each patient’s clinical problem</td>
</tr>
<tr>
<td>PC2</td>
<td>Develops and achieves comprehensive management plan for each patient</td>
</tr>
<tr>
<td>PC3</td>
<td>Manages patients with progressive responsibility and independence</td>
</tr>
<tr>
<td>PC4</td>
<td>Skill in performing procedures</td>
</tr>
<tr>
<td>PC5</td>
<td>Requests and provides consultative care</td>
</tr>
<tr>
<td>PD</td>
<td>Program Director</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS (cont.)

PROF1  Has professional and respectful interactions with patients, caregivers and members of the interprofessional team

PROF2  Accepts responsibility and follows through on tasks

PROF3  Responds to each patient’s unique characteristics and needs

PROF4  Exhibits integrity and ethical behavior in professional conduct

QI     Quality Improvement

RUMC   Rush University Medical Center

SBP1   Works effectively in an interprofessional team

SBP2   Recognizes system error and advocates for system improvement

SBP3   Identifies forces that impact the cost of health care, advocates for, and practices cost-effective care

SBP4   Transitions patients effectively within and across health care delivery systems
SUMMARY

Clinical competency committees (CCCs) play a critical role in determining the developmental progress of residents. This study examines how CCCs synthesize assessment data to make judgments about residents’ clinical performance. Beginning in 2014, CCC faculty completed questionnaires about the CCC process after each six-month reporting period to the Accreditation Council of Graduate Medical Education (ACGME). Subsequently, a focus group interview was conducted. Qualitative data was analyzed using the constant comparative method, consistent with a grounded theory approach. Milestones ratings were analyzed for two ACGME reporting cycles (n=102 categorical residents).

Seven CCC members provided qualitative data for four ACGME reporting cycles. Three major themes were identified: 1) “data-input issues” (i.e., problems with the data or lack thereof); 2) “CCC data processing issues” (i.e., factors influencing ratings and decision-making processes) and 3) “data output issues” (i.e., how CCC generated milestones ratings are used). CCC members weighted rotation evaluation scores highest (weight=37%), followed by comments (weight=27%), and personal experience with residents (weight=14%) to make judgments about learner milestone levels. Generally, milestone levels improved significantly over time for second- and third-year residents; however, for first-year residents, milestone levels during the second reporting cycle were significantly lower. Milestones that received the most “not able to assess ratings” were PC4, PC5, SBP2, PBLI2 and PROF3.

Identifying factors that affect assessment data at all stages of the CCC process can assist in improving assessment systems and in supporting faculty development for CCCs.
In addition, recognizing challenges in synthesizing first- and second-hand assessment data is an important step toward understanding the CCC decision-making process.
I. INTRODUCTION

A. Background

In July 2013, the Accreditation Council for Graduate Medical Education (ACGME) mandated the Next Accreditation System (NAS) (Nasca et al., 2012), which requires residency programs to develop Clinical Competency Committees (CCCs) whose primary goal is to rate each resident biannually on their developmental progress toward competency in their specialty-specific milestones (ACGME, 2009). As part of this process, CCCs must analyze different types of assessment data collected in the context of a variety of resident experiences and make informed judgments about each resident’s clinical performance. CCCs also make recommendations to program directors (PDs) regarding promotion and remediation (ACGME, 2009).

B. Statement of the Problem

Although these newly formed CCCs play a critical role in the assessment of residents, there is limited information in the literature about how CCCs actually perform their work. Although most literature about CCCs has been descriptive, with a focus on developing CCCs (Promes and Wagner 2014, Ketteler et al., 2014, Colbert et al., 2015), there have been attempts to describe the role of CCCs and how assessment data can be organized to facilitate the decision making process. Hauer and colleagues defined the goal of CCCs from PDs’ perspectives as
being either that of “problem identification” (focused on identifying the struggling learner) or “developmental” (e.g., fostering a “learning trajectory” for all residents) (Hauer et al., 2015), with most CCCs having difficulty with the latter. How CCCs perceive their role may impact what types of assessment data they choose to review and how they make decisions about residents’ clinical performance. Friedman and colleagues describe the process of developing a resident dashboard, consisting of different types of assessment data, to assist in organizing and delivering data to CCCs (Friedman et al., 2016). Each data source (e.g., faculty evaluations, peer evaluations) is assigned a “weighted” score, which impacts CCC decision-making about milestones ratings. Although the ACGME has recently provided guidance on CCC policies and procedures (Andolsek et al., 2015), residency programs still have yet to determine exactly how CCCs should make decisions, and as such, a deeper understanding of how CCC members use assessment data to make collective decisions on learners is needed.

C. **Purpose of the study**

This study is designed to investigate how CCCs collect and process assessment data. We examined CCC members’ perspectives on their use of different assessment data and factors that influence their decision-making processes. Both qualitative and quantitative data were used in a mixed-methods approach to address the following questions:
1. **Quantitative:**
   
   a. What “relative weight” do CCC faculty members associate with different types of assessment data?
   
   b. Which reporting milestones do CCC faculty have difficulty assessing and how does this vary by level of training?

2. **Qualitative:**
   
   a. What perspectives do CCC faculty members have about the relative importance of different assessment data?
   
   b. What perspectives do CCC faculty members have about decision-making processes and their role in the CCC assessment system?
II. METHODS

A. Setting and Participants

All 16 CCC members (excluding AE) in the Department of Internal Medicine at Rush University Medical Center (RUMC) in Chicago, Illinois were invited to participate in this study. Seven (7) CCC faculty members (out of 12 CCC faculty; 4 were chief residents) chose to participate in the study. Chief residents were excluded in this analysis.

All participants had been in practice for at least 7 years (range 7-42 years). They represented three broad specialty areas (four hospitalists, three ambulatory care physicians and one subspecialist). There were six assistant professors and one 1 professor. Six of the seven participants reported some training in assessment (e.g., workshop attendance at RUMC, regional and national conferences). Bias analyses were conducted to check for possible differences in CCC member characteristics between the sampled (7 members) and the full CCC (12 eligible). Results showed there were no significant differences in CCC characteristics, including gender, faculty rank, years of experience, areas of specialty, and prior training in assessment.
B. **RUMC CCC Structure and Process**

The RUMC CCC is comprised of four subcommittees (preliminary year and postgraduate year [PGY] 1, 2 and 3). Support for using subcommittees for CCCs in large programs (i.e., over 75 residents) can be found in the ACGME CCC guidebook (Andolsek et al., 2015). Each subcommittee consists of an associate program director, a chief resident who has completed their training at RUMC and two core faculty members. These subcommittees follow a group of residents during all three years of their residency training. During each reporting cycle, each subcommittee member is assigned a group of seven-nine residents to rate. One CCC faculty member rates each resident. The subcommittee members use performance data to rate each resident on the 22 internal medicine (IM) reporting milestones. Each subcommittee meets to discuss all residents in their assigned cohort. During this process, milestones ratings for each resident are discussed and low-performing residents (i.e., learners with multiple negative comments, professionalism issues, or low milestones scores in comparison to their peers) are identified. Subsequently, the CCC meets, and each subcommittee chairperson presents these low-performing residents to the entire CCC for discussion. The learners being rated included all preliminary and categorical residents ($n=115$); however, this study focuses primarily on categorical residents ($n=102$) i.e. learners with longitudinal assessment data.
C. **Study Design: Data Analysis**

1. **Quantitative Data:** On items Two, Three and Four of the questionnaire, we asked participants to weigh (assign a percentage value) the relative importance of different types of CCC data, i.e., including end-of-rotation evaluation scores (milestones summary data generated by our residency management system, MedHub), comments on end-of-rotation evaluation forms, first-hand knowledge based on working directly with a trainee (personal experience), information obtained informally about a trainee from colleagues, CCC large group discussions and CCC subcommittee discussions.

   We also analyzed milestones ratings generated by all of our CCC members from two of the earlier reporting cycles (December 2014 and June 2015). Descriptive statistics were calculated to examine trends on each of the 22 milestones and further categorized by level of postgraduate training levels. This analysis included frequency of the use of “Not able to assess” (NA) in the rating process.

2. **Qualitative Data:** The qualitative data consisted of (1) responses on the questionnaire given after each reporting period to the ACGME, (2) follow-up questions via member checking, and (3) a semi-structured focus group interview.
After each reporting cycle to the ACGME, (6/2014, 12/2014, 6/2015, 12/2015), participants were asked to complete a questionnaire regarding their perspectives about rating residents on their achievement of the milestones (see Appendix A).

Schon’s reflective practice (Schon, 1984) was used as a conceptual framework in deciding how to ask participants about their work. CCC faculty members used “reflection on action” (i.e., to reflect about their practice) to complete the questionnaire and participate in the focus group interview. We also used the conceptual framework developed by Hauer and colleagues, “problem identification” versus “developmental model”, to revise items on the questionnaires. Participants completed and submitted questionnaires electronically (Hauer et al., 2015). Responses were not anonymous, to aid in member checking. The questions were designed to elicit responses about the use of the milestones as an assessment framework, the role of different types of data in the CCC’s decision making process, and to provide insights about the strengths, weakness, and opportunities for improvement in the CCC assessment system.

The initial questionnaire items were developed based on a review of the CCC literature, including identified knowledge gaps, questions that arose during actual CCC meetings and an initial analysis of the milestones data. Subsequently, each version of the questionnaire contained a few modifications based on the analysis of both the qualitative and quantitative data from the previous reporting cycle. Study participants were surveyed after the CCC meetings to avoid
potentially influencing their milestones ratings and to capture reflections on the actual ratings assigned by the CCC faculty.

Using phrases as the unit of analysis, the investigators (AE and EB) independently read and identified themes in the data from the first questionnaire. Themes were identified through an iterative process, using the constant comparative method of analysis associated with grounded theory approaches (Charmaz, 2014). After the first reporting period, an external reviewer was asked to review the data analysis. Subsequently, the two investigators coded data from the additional questionnaires, using the agreed upon coding guide. They met multiple times to discuss and clarify the themes. Member checking was performed, as appropriate, to improve the trustworthiness of the data analysis. After the third cycle, no new themes were identified and after the fourth reporting period, a fourth investigator reviewed the data analysis.

To determine if thematic saturation had been reached, one focus group interview (with all participants) was conducted (see Appendix B for the focus group guide). The focus group questions were developed based on an analysis of the responses to items on previously administered questionnaires. The format for the session was based on recommendations by Stalmeijer and colleagues (Stalmeijer et al., 2014). A script was developed for the focus group interview. The focus group was 90 minutes, with the first 30 minutes spent completing the post-reporting period questionnaire, which primed the participants for the subsequent 60-minute
Six CCC members participated in the interview. The interview was audiotaped without identifiers and transcribed. Upon review, it was concluded that saturation of themes had been achieved.

The Institutional Review Boards of Rush University and the University of Illinois at Chicago approved this study.
III. RESULTS

A. Quantitative Data

Weights for different assessments: CCC faculty participants were asked to weigh factors contributing to their milestones ratings. Figure 2 illustrates the weighted values for each type of data and also the variability associated with each mean value (95% confidence intervals). Faculty weighted rotation evaluation scores highest (weight=37%), followed by comments (weight=27%), and personal experience (weight=14%).

Trends in Milestone ratings: Mean milestone ratings, and standard deviations for residents at each PGY level and across two reporting periods, were calculated (see Table 1 and Figure 2). At the PGY-1 level, mean scores were lower during the second reporting period when compared to the first. This trend was not present at the PGY-2 and PGY-3 levels. Milestones with the highest frequency of “Not able to assess” ratings were PC4, PC5, SBP2, PBLI2 and PROF3 (see Table 2).

B. Qualitative Data

Nine initial themes were identified and collapsed into three major themes and associated subthemes. The three major themes and subthemes are presented in Table 3 and Figure 3. All three major themes were based on “data-related factors”: (1) “data input issues” (i.e., factors
related to the data input to the CCC, (2) “CCC data processing issues” (i.e., decision making processes), and (3) “data output issues” (i.e., how CCC-generated milestones data is used).

1. **Data input**

Data Quantity: Participants identified the need for additional data in order to accurately rate resident performance on the milestones. For instance, regarding procedural skills, one participant reported, “We lack data and we lack consensus about which procedures should be included.”

Data Quality: There was concern about not having a shared mental model with respect to expected performance criteria, with a participant commenting, “so much of it comes back to the discomfort of not quite knowing what criteria the evaluators are using.” In addition, although faculty comments on end-of-rotation evaluations were generally thought to be useful, they were sometimes viewed as being too vague. Being on the CCC also provided participants with an appreciation for the complexity of the assessment system. As one participant commented, “I think it’s much more difficult than I thought in terms of collecting data, quality data, reliable data. I had no idea. I had no idea how much work went into creating the individual evaluations that are used in every rotation and how painstaking that must have been.”
2. **CCC data processing issues**

Trustworthiness of first-hand versus second-hand data: One participant described the struggle in determining the value of first-hand knowledge, in relation to the issue of fairness, commenting, "*My understanding of the committee process is that we are to collect, correlate and synthesize the evaluation data. Clearly our own personal knowledge of a resident can't be excluded from this process, but the design of the CCC and ABIM (ACGME) mandate suggests that we are supposed to gather information as neutral reporters rather than as candidate advocates. We should acknowledge our own personal observations and opinions, and then work to consider them as one bit of input among many...if a resident accumulates six or seven evaluations over six months, our own personal knowledge surely ought to weigh no more than one of those evaluations, or less than 15%.*"

Participants valued second hand information as confirmatory data for their own initial judgments. As one participant commented, "*Helps to get information from others who have worked with the resident.*” Participants expressed awareness of the difficulty of trying to use data provided by a colleague. As one participant commented, "*You’re actually in essence evaluating the resident as well, as you evaluate basically... because you don’t know how that evaluator... you don’t know what they are thinking. It’s not like you have the luxury of immediately asking the evaluator, “What were you evaluating?”*"
Initially, participants relied heavily on qualitative data (e.g., faculty comments) over quantitative data (e.g., rotation evaluation scores). Participants commented, “Sometimes the comments gave me more concrete data about certain milestones than the numeric evaluations did.” “Comments help when a resident might be borderline between the well above, average, or not meeting expectations. Gives a broader understanding of the resident in a way the points/numbers can’t quite capture.” However, more recently, participants discussed their concern about fairness and being unsure about using comments when they do not have a shared mental model with the faculty providing the comments. A participant commented, “I was actually surprised because when I answered question two, I put (rated) comments very low, because I felt that if I don’t use this milestone summary score, I’m not being fair to all the residents and if I’m just basing my evaluations on comments and I’m not sure who’s commenting on this...I always use comments but now that I am part of this committee, I’m trying not to. I’m trying to be fair and transparent as much as I can...I actually only put 10% (value placed on comments) but I always have used comments in the past, but right now, for me it was harder because I’m trying to be equal and that’s why I use those milestones (summary score).”

Qualitative versus Quantitative data: Participants struggled with how much value to attach to quantitative versus qualitative data. A participant commented, “…I would put down the summary scores (MedHub generated milestones score reports) considerably more than the
comments because many of the evaluations don’t have comments or many of the comments are of the nature of “great resident” and that doesn’t help. I went with equal weighting for mine between those two as well because the comments are helpful when it swings one way or the other but it is again, very vanilla a lot of times. In addition, I fall into the trap of I don’t know the commenter as well.”

The milestones framework: Our participants were both complimentary and critical of the milestones framework.

Positive insights: Participants commented, “…Able to monitor and evaluate trajectory of growth- including plateaus and unlinked to year of training for each resident…” “It standardizes the process and the numeric scale makes it easier to compare residents.”

Critique of milestones scale: There were multiple comments on the difficulty encountered when using the milestones rating scale, including range restriction based on the narrative description along the scale (especially when rating interns) and not being able to distinguish between the nine levels of performance. Participants commented, “Forces you to categorize residents into these specific, fairly rigid, performance levels for each individual reporting milestone.” “The name ‘milestones’ is a malapropism, as it implies we are reporting on completion of tasks and definitive acquisition of skills; in actuality we report degrees of competence and increasing sophistication of thinking and acting. I think the present reporting system implies that we know
more than we actually do about our residents’ competencies.” “For interns there are only two possible grades.” “When completing the milestones evaluations, the answers (narrative anchors) are too long and somewhat confusing- there are nine potential categories within which a resident may fall, which seems like too many...”

3. **Data output from the CCC**

Clarifying the role of the CCC: Participants expressed the need for guidance from leadership regarding the role of the CCC and how CCC-generated data will be used. A participant commented, “**We should clarify the purpose of the milestones evaluations and what use will be made of them. It appears the (ABIM?) wants them to make sure no defective residents graduate from the program. This is a valid purpose, but obviously residents with substandard performance should be identified and intervened upon before six months (or a year) have gone by. Further, if this is the only purpose, then we could probably simplify the process a great deal and rely on other ways to identify the stars.**” It is possible that uncertainty about these factors plays a role in their ratings.

Giving Feedback to Residents: Feedback to residents was seen as an important responsibility of the CCC, but participants expressed concern about not having a standardized process for providing feedback, as well as a lack of formal remediation programs for those residents with identified deficiencies.
Identifying struggling residents: The issue of who is responsible for identifying struggling residents, and the time frame for this process within a yearly cycle, was discussed. Is the CCC expected to identify struggling residents from the outset or is the role of the CCC to just confirm the list of struggling residents already identified by program? Participants commented, “I know it’s ACGME mandated but I often feel like (the CCC) is a superfluous committee, only because I feel like we identify residents that have been identified already by the program director or the chiefs...” “The only time it seems like we bring (new) information is (when) sometimes a CCC member will say, “This resident had the most impressive, outstanding comments” and the chiefs will fall backwards in their chair and say “We have had terrible reviews about this person all year.” That’s always interesting. I feel like a lot of information that comes forth has already been brought forth.”

Curriculum and Assessment issues: Participants recognized that the CCC uncovers curriculum and assessment issues that need to be addressed. However, it was unclear as to who should be responsible for addressing these issues -the CCC, the program evaluation committee, or another group.

Recommendations about CCC structure and function: Participants had multiple recommendations to make regarding the structure and functions of the CCC. These recommendations included assigning CCC faculty to residents that they have directly worked
with and continuing to have CCC members “follow” a particular cohort of residents throughout their training. They recommended continuing to work on getting the “right data” to the CCC e.g., from learning experiences that would help provide assessment data for those milestones with which we are currently having difficulty. They proposed providing core faculty with feedback about their ratings and suggested inviting core faculty to CCC meetings in order to explain their evaluations.

Quality improvement for the CCC itself: Comments about quality improvement for the CCC included both faculty development needs and insights about the membership and structure of the committee.

Faculty development for CCC faculty: On the initial questionnaire, when asked what participants would do differently during the next reporting cycle, or what suggestions they had for the residency program and department leadership, participants had difficulty expressing their faculty development needs e.g. “Not sure how to do it better...Too soon to consider making changes.” However, in subsequent questionnaires and in the focus group, they advocated for faculty development, not only for themselves but also for others in the assessment system (e.g., core faculty, subspecialists). In addition, they recognized the importance of residents understanding the CCC process. They appreciated the complexity of assessment and the benefit of learning from each other, including asking to review each other’s ratings in order to have a
better understanding of the group’s process and values (Hauer et al., 2016). They also asked for guidelines and best practices. They wanted feedback on their milestones ratings and wanted to see their scores compared to those of the other CCC members. A participant commented, “The question about would I want to see my score is that I think it would be good to see everyone’s scores and see how they trend.” Participants described not understanding the behind-the-scenes scaffolding of the assessment system, i.e., not knowing the connection between items on our end of rotation assessment forms and the reporting milestones i.e. “what feeds into the milestones?”

Review of CCC membership: The strengths of our CCC were identified as being a collaborative working group with multiple member input, some members having first-hand knowledge of their assigned residents and some diversity within the membership, including subspecialists and faculty from both the inpatient and outpatient settings.

CCC structure: Participants felt that it was important to review the purposes of the subcommittee structure and to compare practices amongst them
IV. DISCUSSION

The results of this mixed methods study revealed not only how CCC members use assessment data in assessing resident performance, but also how this data impacts their work and their perceptions about their roles. Three major themes were identified (data input, CCC data processing, and data output), highlighting important concerns for CCC faculty in the process of conducting their work.

In our study, participants relied primarily on end-of-rotation evaluation scores and faculty comments in rating residents on the milestones. Interestingly, they did not weigh the value their own first-hand knowledge of a resident’s clinical skills as highly other data. We also found that having more first-hand knowledge about particular residents did not necessarily result in higher ratings for these residents. This finding may be a reflection of participants' concern about issues of fairness mentioned earlier, a lack of confidence in their direct observations of residents, recall bias, or other factors. Also, there was more variability in the weighted score for mean evaluations, comments and personal (first hand) knowledge compared to other types of data. The qualitative data support these findings.

We also identified an additional layer of complexity in the CCC role. Some CCC members may not have experience using second-hand data as the only information source to render judgments about trainee performance. On the other hand, participants with positions such
as being a course director or clerkship director may have some experience with not having the benefit of their own direct observations when such decisions need to be made. Thus, the CCC role requires faculty to utilize assessment data from secondary sources i.e. “meta-assessment”.

As CCCs attempt the difficult task of both identifying struggling learners and ensuring that all residents are on the appropriate trajectory towards competence as outlined by Hauer and colleagues (Hauer et al., 2015), the importance of CCC faculty having first-hand knowledge of residents’ clinical skills to make judgments requires further study. Perhaps this concern is unique to moderate and large sized training programs, but it also raises the issue of fairness in the assessment process. Should residents always be assigned to CCC faculty raters who have first-hand knowledge of their skills or should CCC faculty always be “blinded” to their assigned residents?

A. **Context**

Rather than conceptualizing different types of assessment data (e.g. direct observation, chart audits) or types of assessment tools, we chose to focus on data sources that our study participants were most familiar with (e.g., end-of-rotation comments), and methods that were recently developed as a result of NAS (i.e., milestones summary score, CCC meetings).

Themes that were identified in the analysis of the qualitative data, such as issues of uncertainty about the role of the CCC, trying to merge first-hand knowledge of a resident with other
important sources of information, and the need for the CCC to be alerted to performance
problems in real time, have been discussed by Hauer and colleagues (Hauer et al., 2015). Despite
familiarity with the ACGME’s description of CCCs, our participants repeatedly asked for
clarification of the role of the CCC and the impact of their ratings. Having a better
understanding of their role may impact the type of assessment data they request, and how they
synthesize it, make judgments and determine how these judgments affect key stakeholders and
drive quality improvement in the larger assessment system. Having a clear purpose will also
help to determine how best to structure CCCs and support their efforts. Although study
participants valued their first-hand knowledge of a resident’s clinical skills, they looked to
secondary data sources to provide fairness in the decision making process.

Our quantitative data revealed that our CCC faculty had difficulty assessing performance
for some milestones (e.g., PC4, PC5) that have been found to be problematic to assess for others
as well (Hauer et al., 2016b). In addition, we found that our CCC faculty were judicious in their
rating of PGY3 residents, as evidenced by not all graduating residents receiving Level 5 ratings
prior to graduation (Hauer et al., 2016b). The lower scores for PGY1s at the June 2015 reporting
cycle (compared to their December 2014 scores) may reflect the increasing assessment skills of
their CCC raters as the raters attempted to calibrate their expectations for resident performance
with the rating scale anchors. Our study participants specifically mentioned the difficulty of
trying to accurately rate the PGY1s, given the tone of the narrative descriptors towards the left and middle of the rating scale, in the context of allowing for the development of a trajectory of performance improvement. The similarities between the trends in our scores and that which has been reported in the literature lends some credibility to our CCC process. The analysis of qualitative data helped explain some of the trends in the milestones scores.

B. **Strengths and Limitations**

The strengths of this study lie in its longitudinal and mixed methods approach. As part of the study design, we purposefully scheduled the focus group at the midpoint of the study, as we wanted to use the data from the questionnaires and milestones scores to help guide us towards issues that might require more in-depth probing. The collection of data over a prolonged period of time not only provided participants the opportunity to discover their role in the assessment system and their faculty development needs, but also led to a more in-depth discussion during the focus group, which added to the richness of the data. The time factor proved an important issue, as CCCs are still struggling to define their role since the implementation of NAS in 2013. We attempted to validate our qualitative findings through data triangulation (use of questionnaires and focus group interview), auditing (use of external reviewers), and member checking. The use of the quantitative data helped inform our initial development of our questionnaire items and
continued to serve as a checkpoint to determine any new trends in their ratings requiring exploration with qualitative approaches.

There are limitations to this study. First, it was challenging to administer the questionnaires immediately after the study participants performed their milestones assessments, therefore, there may be a recall bias. Second, our questionnaires were not anonymous; however this allowed for member checking when needed. Third, due to the size of our program, only one CCC faculty was assigned to rate each resident, making it challenging to develop a shared mental model and discuss specific milestones ratings. Finally, this study was performed using a single CCC within a single institution. To the extent that institutional context and culture influences the assessment system, it is possible that results may differ in other institutions and specialties.

C. **Implications**

The study findings are important as we aim for transparency in our assessment of residents, provide feedback to other stakeholders, and improve the quality of our CCC. In addition to faculty development workshops on topics such as rater training, and group process, attention should be given to developing a shared mental model of how value is placed on different types of assessment data and the added complexity of their role as not only assessors but “meta-assessors. The three main themes highlighting the factors affecting the quality and
impact of their work signify another potential role of CCCs — that of driving quality improvement efforts for assessment systems.
V. CONCLUSION

This study provides themes for conceptualizing how CCCs value different types of assessment data and the concerns they have at different stages of data syntheses and use (input, CCC processing, and output) in the assessment system. These issues should further be explored in the determination of faculty development content for CCCs. Thus, this study facilitates the generation of ideas and hypotheses for further research. In addition, the concept diagram for data flow through the CCC (Figure 3) may be useful as programs try to address issues raised about how to construct their assessment systems to best serve trainees, CCCs, and ultimately, patients. Future work should address more granular questions about the types of data needed to assess each of the 22 Internal Medicine subcompetencies (Jobst et al., 2015) and how such data can be optimally synthesized to make judgments about resident performance.
<table>
<thead>
<tr>
<th>Subcomp</th>
<th>December 2014</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>June 2015</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PGY-1</td>
<td>PGY-2</td>
<td>PGY-3</td>
<td></td>
<td>PGY-1</td>
<td>PGY-2</td>
<td>PGY-3</td>
<td></td>
<td>PGY-1</td>
<td>PGY-2</td>
<td>PGY-3</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
<td>Mean (SD)</td>
<td></td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>PC1</td>
<td>33</td>
<td>2.82 (.24)</td>
<td>34</td>
<td>3.19 (.25)</td>
<td>32</td>
<td>3.89 (.30)</td>
<td>33</td>
<td>2.71 (.28)</td>
<td>34</td>
<td>3.49 (.57)</td>
<td>33</td>
</tr>
<tr>
<td>PC2</td>
<td>32</td>
<td>2.78 (.28)</td>
<td>34</td>
<td>3.24 (.25)</td>
<td>32</td>
<td>3.94 (.30)</td>
<td>31</td>
<td>2.77 (.36)</td>
<td>34</td>
<td>3.50 (.58)</td>
<td>33</td>
</tr>
<tr>
<td>PC3</td>
<td>24</td>
<td>2.71 (.29)</td>
<td>30</td>
<td>2.98 (.36)</td>
<td>32</td>
<td>3.98 (.52)</td>
<td>32</td>
<td>2.77 (.36)</td>
<td>30</td>
<td>3.33 (.59)</td>
<td>32</td>
</tr>
<tr>
<td>PC4</td>
<td>12</td>
<td>2.58 (.19)</td>
<td>14</td>
<td>2.86 (.36)</td>
<td>25</td>
<td>3.86 (.42)</td>
<td>11</td>
<td>2.59 (.30)</td>
<td>19</td>
<td>3.21 (.51)</td>
<td>29</td>
</tr>
<tr>
<td>PC5</td>
<td>10</td>
<td>2.50 (.00)</td>
<td>10</td>
<td>2.95 (.37)</td>
<td>29</td>
<td>4.00 (.38)</td>
<td>10</td>
<td>2.50 (.00)</td>
<td>17</td>
<td>3.12 (.22)</td>
<td>29</td>
</tr>
<tr>
<td>MK1</td>
<td>33</td>
<td>2.88 (.28)</td>
<td>34</td>
<td>3.18 (.30)</td>
<td>32</td>
<td>3.83 (.33)</td>
<td>33</td>
<td>2.70 (.30)</td>
<td>34</td>
<td>3.44 (.55)</td>
<td>33</td>
</tr>
<tr>
<td>MK2</td>
<td>33</td>
<td>2.83 (.24)</td>
<td>34</td>
<td>3.25 (.28)</td>
<td>32</td>
<td>3.80 (.38)</td>
<td>33</td>
<td>2.70 (.28)</td>
<td>34</td>
<td>3.50 (.63)</td>
<td>33</td>
</tr>
<tr>
<td>SBP1</td>
<td>33</td>
<td>2.79 (.28)</td>
<td>34</td>
<td>3.13 (.22)</td>
<td>32</td>
<td>3.94 (.33)</td>
<td>33</td>
<td>2.67 (.32)</td>
<td>34</td>
<td>3.60 (.69)</td>
<td>33</td>
</tr>
<tr>
<td>SBP2</td>
<td>23</td>
<td>2.65 (.24)</td>
<td>14</td>
<td>3.04 (.41)</td>
<td>26</td>
<td>3.98 (.41)</td>
<td>27</td>
<td>2.59 (.31)</td>
<td>18</td>
<td>3.11 (.32)</td>
<td>27</td>
</tr>
<tr>
<td>SBP3</td>
<td>33</td>
<td>2.65 (.23)</td>
<td>31</td>
<td>3.00 (.22)</td>
<td>32</td>
<td>3.94 (.40)</td>
<td>33</td>
<td>2.61 (.30)</td>
<td>33</td>
<td>3.42 (.56)</td>
<td>33</td>
</tr>
<tr>
<td>SBP4</td>
<td>30</td>
<td>2.63 (.22)</td>
<td>32</td>
<td>3.03 (.25)</td>
<td>32</td>
<td>3.92 (.38)</td>
<td>33</td>
<td>2.61 (.37)</td>
<td>32</td>
<td>3.42 (.61)</td>
<td>33</td>
</tr>
<tr>
<td>PBL11</td>
<td>33</td>
<td>2.70 (.28)</td>
<td>31</td>
<td>3.08 (.23)</td>
<td>32</td>
<td>3.95 (.43)</td>
<td>33</td>
<td>2.62 (.31)</td>
<td>33</td>
<td>3.47 (.59)</td>
<td>33</td>
</tr>
<tr>
<td>PBL12</td>
<td>10</td>
<td>2.50 (.00)</td>
<td>12</td>
<td>3.04 (.33)</td>
<td>26</td>
<td>4.04 (.51)</td>
<td>11</td>
<td>2.45 (.15)</td>
<td>18</td>
<td>3.14 (.29)</td>
<td>22</td>
</tr>
<tr>
<td>PBL13</td>
<td>33</td>
<td>2.77 (.28)</td>
<td>31</td>
<td>3.06 (.25)</td>
<td>32</td>
<td>3.95 (.45)</td>
<td>33</td>
<td>2.62 (.31)</td>
<td>30</td>
<td>3.40 (.58)</td>
<td>33</td>
</tr>
<tr>
<td>PBL14</td>
<td>33</td>
<td>2.86 (.29)</td>
<td>34</td>
<td>3.32 (.27)</td>
<td>32</td>
<td>3.92 (.34)</td>
<td>33</td>
<td>2.68 (.35)</td>
<td>34</td>
<td>3.62 (.63)</td>
<td>33</td>
</tr>
<tr>
<td>PROF1</td>
<td>33</td>
<td>2.85 (.23)</td>
<td>34</td>
<td>3.22 (.28)</td>
<td>32</td>
<td>3.94 (.28)</td>
<td>33</td>
<td>2.74 (.25)</td>
<td>33</td>
<td>3.61 (.70)</td>
<td>33</td>
</tr>
<tr>
<td>PROF2</td>
<td>33</td>
<td>2.83 (.24)</td>
<td>34</td>
<td>3.22 (.25)</td>
<td>32</td>
<td>3.95 (.34)</td>
<td>33</td>
<td>2.74 (.28)</td>
<td>34</td>
<td>3.62 (.71)</td>
<td>33</td>
</tr>
<tr>
<td>PROF3</td>
<td>18</td>
<td>2.86 (.38)</td>
<td>9</td>
<td>3.39 (.42)</td>
<td>25</td>
<td>3.92 (.34)</td>
<td>27</td>
<td>2.76 (.40)</td>
<td>17</td>
<td>3.29 (.36)</td>
<td>27</td>
</tr>
<tr>
<td>PROF4</td>
<td>33</td>
<td>2.76 (.31)</td>
<td>32</td>
<td>3.11 (.33)</td>
<td>33</td>
<td>4.14 (.39)</td>
<td>33</td>
<td>2.67 (.27)</td>
<td>34</td>
<td>3.50 (.63)</td>
<td>33</td>
</tr>
<tr>
<td>ICS1</td>
<td>33</td>
<td>2.79 (.25)</td>
<td>32</td>
<td>3.06 (.28)</td>
<td>32</td>
<td>4.03 (.40)</td>
<td>33</td>
<td>2.61 (.32)</td>
<td>33</td>
<td>3.45 (.63)</td>
<td>33</td>
</tr>
<tr>
<td>ICS2</td>
<td>33</td>
<td>2.86 (.26)</td>
<td>34</td>
<td>3.26 (.25)</td>
<td>32</td>
<td>3.97 (.31)</td>
<td>33</td>
<td>2.68 (.27)</td>
<td>34</td>
<td>3.62 (.73)</td>
<td>33</td>
</tr>
<tr>
<td>ICS3</td>
<td>33</td>
<td>2.88 (.28)</td>
<td>34</td>
<td>3.28 (.25)</td>
<td>32</td>
<td>3.94 (.28)</td>
<td>33</td>
<td>2.70 (.28)</td>
<td>34</td>
<td>3.56 (.74)</td>
<td>33</td>
</tr>
</tbody>
</table>

Note: Subcompetencies (Iobst et al., 2015): “PC1=Gathers and synthesizes essential and accurate information to define each patient’s clinical problem(s), PC2=Develops and achieves comprehensive management plan for each patient, PC3=Manages patients with progressive responsibility and independence; PC4=Skill in performing procedures, PC5=Requests and provides consultative care, MK1=Clinical knowledge, MK2=Knowledge of diagnostic testing and procedures, SBP1= Works effectively within an interprofessional team, SBP2 Recognizes system error and advocates for system improvement, SBP3=Identifies forces that impact the cost of health care, advocates for, and practices cost-effective care, SBP4=Transitions patients effectively within and across health delivery systems, PBL1=Monitors practice with a goal for improvement, PBL2= Learns and improves via performance audit, PBL3= Learns and improves via feedback, PBL4= Learns and improves at the point of care, PROF1= Has professional and respectful interactions with patients, caregivers and members of the interprofessional team, PROF2= Accepts responsibility and follows through on tasks, PROF3= Responds to each patient’s unique characteristics and needs, PROF4= Exhibits integrity and ethical behavior in professional conduct, ICS1= Communicates effectively with patient and caregivers, ICS2=Communicates effectively in interprofessional teams, ICS3=Appropriate utilization and completion of health records”
TABLE II. FREQUENCY OF “NOT ABLE TO ASSESS (NA) RATINGS OVER 2 REPORTING CYCLES - DEC 2014 AND JUN 2015

<table>
<thead>
<tr>
<th>Subcompetency</th>
<th>DEC 2014</th>
<th></th>
<th></th>
<th>Total</th>
<th>DEC 2014</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PGY1</td>
<td>PGY2</td>
<td>PGY3</td>
<td>Total</td>
<td>PGY1</td>
<td>PGY2</td>
<td>PGY3</td>
<td>Total</td>
</tr>
<tr>
<td>PC1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PC2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>PC3</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>13</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>PC4</td>
<td>21</td>
<td>20</td>
<td>7</td>
<td>48</td>
<td>22</td>
<td>15</td>
<td>4</td>
<td>41</td>
</tr>
<tr>
<td>PC5</td>
<td>23</td>
<td>24</td>
<td>3</td>
<td>50</td>
<td>23</td>
<td>17</td>
<td>4</td>
<td>44</td>
</tr>
<tr>
<td>MK1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MK2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SBP1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SBP2</td>
<td>10</td>
<td>20</td>
<td>6</td>
<td>36</td>
<td>6</td>
<td>16</td>
<td>6</td>
<td>28</td>
</tr>
<tr>
<td>SBP3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SBP4</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>PBLI1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PBLI2</td>
<td>23</td>
<td>22</td>
<td>6</td>
<td>51</td>
<td>22</td>
<td>16</td>
<td>11</td>
<td>49</td>
</tr>
<tr>
<td>PBLI3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>PBLI4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PROF1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PROF2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PROF3</td>
<td>15</td>
<td>25</td>
<td>7</td>
<td>47</td>
<td>6</td>
<td>17</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>PROF4</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ICS1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ICS2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ICS3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>105</td>
<td>130</td>
<td>29</td>
<td>264</td>
<td>82</td>
<td>95</td>
<td>32</td>
<td>209</td>
</tr>
<tr>
<td>Themes</td>
<td>Subthemes</td>
<td>Details</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-----------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 1. Data Input to the CCC                    | (1) Problems with data quantity                     | • Not enough data especially for marginal/extreme residents and certain milestones  
• Different formats of data (paper + electronic)  
• Data not helping to distinguish between performance levels  
• Not feeling confident about the accuracy of the data |
|                                             | (2) Problems with data quality                      | • First-hand knowledge versus second-hand information (e.g. faculty comments, CCC member input, informal data)  
• Not knowing the evaluator  
• Lack of a shared mental model |
|                                             | (1) Trustworthiness of first-hand versus second-hand data | • Normative vs. criterion standard (comparing residents to previous residents)  
• Trying to avoid “gestalt”  
• Difficult to translate faculty evaluations into a rating on the milestones scale  
• Mismatch between milestones rating and comments: Poor quality of some comments, |
|                                             | (2) Qualitative versus quantitative data             | • Can be used to monitor a resident’s competency trajectory  
• Easy to compare a given resident to their peers  
• Helps identify curriculum and assessment gaps  
• Seems like a more intuitive assessment framework  
• Standardizes the assessment process across residents (for comparisons), locations, provides uniform language  
• Helps align learning objectives with assessments  
• Creates aspirational goals  
• Fosters more comprehensive assessment  
• More behavioral (concrete behaviors) |
| 2. CCC Data Processing Issues               | (3) Positive insights gained about the milestones framework | • Does not accurately capture residents’ competence  
• Too many milestones/too time consuming  
• Difficult to understand and assess  
• Does not provide specific rating criteria e.g., what procedures? How many of each?  
• Problems with the milestones rating scale  
• Range restriction- especially for interns  
• Hard to distinguish between residents in the “middle” |
### TABLE III. MAJOR THEMES AND SUBTHEMES FROM CCC MEMBERS (cont.)

<table>
<thead>
<tr>
<th>3. Data Output from the CCC</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(1) Help clarify the role of the CCC</strong></td>
<td></td>
</tr>
<tr>
<td>• Unable to distinguish between nine levels of performance</td>
<td></td>
</tr>
<tr>
<td>• Identifying struggling resident vs. CCC just affirming identification of struggling residents already identified by program</td>
<td></td>
</tr>
<tr>
<td>• CCC process needs to be in “real-time” - (not retrospective)</td>
<td></td>
</tr>
<tr>
<td>• Fulfill ACGME requirements</td>
<td></td>
</tr>
<tr>
<td>• System review (30,000 feet)</td>
<td></td>
</tr>
<tr>
<td>• Fix issues identified in the curriculum/assessment system</td>
<td></td>
</tr>
<tr>
<td><strong>(2) Feedback to residents/Documentation for ACGME</strong></td>
<td></td>
</tr>
<tr>
<td>• No standardized process for providing feedback to residents</td>
<td></td>
</tr>
<tr>
<td>• Unclear who is responsible for identifying struggling residents</td>
<td></td>
</tr>
<tr>
<td><strong>(3) Develop expectations of the Department of Medicine</strong></td>
<td></td>
</tr>
<tr>
<td>• Continue to support and recognize the CCC</td>
<td></td>
</tr>
<tr>
<td>• Clarify how milestones data will be used</td>
<td></td>
</tr>
<tr>
<td>• Continue using MedHub</td>
<td></td>
</tr>
<tr>
<td>• Define role of the CCC</td>
<td></td>
</tr>
<tr>
<td>• Allot specific/consistent meeting time</td>
<td></td>
</tr>
<tr>
<td><strong>(4) Feedback to core faculty</strong></td>
<td></td>
</tr>
<tr>
<td>• Provide core faculty with feedback about their ratings</td>
<td></td>
</tr>
<tr>
<td>• Invite them to CCC meetings to explain ratings</td>
<td></td>
</tr>
<tr>
<td><strong>(5) Recommendations to the residency program</strong></td>
<td></td>
</tr>
<tr>
<td>• Continue to modify the current assessments</td>
<td></td>
</tr>
<tr>
<td>• Create pathway for evaluations from unique rotations to the CCC</td>
<td></td>
</tr>
<tr>
<td>• Assign CCC members only the files of residents they have worked with</td>
<td></td>
</tr>
<tr>
<td>• Assign CCC members to particular cohorts of residents “to follow”</td>
<td></td>
</tr>
<tr>
<td>• MedHub is an adequate management system for housing this assessment data and generating reports</td>
<td></td>
</tr>
<tr>
<td>• Unclear remediation options once problems are identified</td>
<td></td>
</tr>
<tr>
<td><strong>(6) Quality improvement for the CCC itself</strong></td>
<td></td>
</tr>
<tr>
<td>• Faculty development for CCC faculty e.g. importance of group process and creating a shared mental model (see each other’s ratings for trends), develop best practices/guidelines, know what (assessment tools/data) “feeds into the milestones”</td>
<td></td>
</tr>
<tr>
<td>• Diversity in the membership</td>
<td></td>
</tr>
<tr>
<td>• Review committee structure- goals of full committee vs. subcommittees</td>
<td></td>
</tr>
<tr>
<td>• CCC work constraints – Lack of time</td>
<td></td>
</tr>
<tr>
<td>• CCC members with first-hand knowledge of residents</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Mean relative weights (%) of different types of assessments\textsuperscript{1-7}: Data aggregated over reporting periods (Mean ± 95% confidence intervals).

Note:

1. “Evaluation” is mean rotation evaluation scores (completed by faculty raters): Mean = 37%, SD = 21%
2. “Comments” is qualitative comments from rotation evaluations: Mean = 27%, SD = 11%
3. “Peers” is information from peers: Mean = 9%, SD = 4%
4. “Personal” is personal experience: Mean = 14%, SD = 11%
5. “Subcommittee” is CCC subcommittee meeting: Mean = 8%, SD = 3%
6. “Whole” is CCC whole group meeting: Mean = 6%, SD = 5%
7. Others included in the survey, but all CCC members reported 0% weight for this category.
Figure 2. Milestones levels by core competency for each postgraduate year levels across two reporting periods: Mean ± 95% confidence intervals.
Figure 3. Factors affecting data flow through the CCC

- Data Input to CCC
  - Data Security
  - Data Quality

- CCC Data Processing
  - Integration and knowledge of policies
  - Data analysis: qualitative & quantitative
  - Use of assessment frameworks

- CCC Data Output
  - Role of CCC
  - How will data be utilized
  - Feedback from staff
  - Accreditation
  - Program evaluation
  - Sustainability and improvement of CCC
  - Faculty development
  - CCC leadership and structure

Impact on Patient Care
APPENDICES

APPENDIX A

CCC Questionnaire- (for reporting period #3, June 2015)

Name:_____________________________________
Date:____________________________________

Thank you for agreeing to participate in this study (ORA 14082705-IRB01). Please complete this questionnaire prior to the start of the interview (focus group). Please answer the questions below to the best of your ability- being as descriptive as possible. If you are unable to answer a question, please state why-e.g. not sure what the question means, new to the CCC and as such have never evaluated residents using the reporting milestones (as might be the case for chief residents), etc.

For questions 1-12, based on the last reporting by our CCC to the ACGME in June 2015:

1) What do you perceive to be the **STRENGTHS** of our CCC? And why?

2) What do you perceive to be the **WEAKNESSES** of our CCC? And why?

3) What do you perceive to be some **OPPORTUNITIES** for our CCC? (e.g. to get involved in faculty development for core faculty, to revise our assessment system, offer feedback to residents about their performance on the milestones). Please explain.

4) What do you perceive to be some of the **THREATS** to the work of our CCC?

5) Please describe if/how you used the milestone score provided by MedHub for each milestone, e.g. PC1-3.2)
6) Most of our assessment data is generated via end-of-rotation evaluation forms. Please describe any other types of data you would like to have in order to assess residents’ performance on the milestones (refer to the attached milestones form if need be).

7) What role (if any) did end-of-rotation evaluation form “comments” play in your decisions?

8) What role (if any) did discussion with other members on your committee play in your decisions?

9) What role (if any) did your own knowledge of a resident (e.g. if you worked with them in the past) play a role in your decisions?

10) What role (if any) did your impressions about a resident based on information from a colleague (or other sources) play a role in your decisions?

11) Based on your responses to questions 7-10 above, please indicate how much (%) each of the following factors contributed to your ratings. The total should add up to 100 e.g.

- Milestones summary score- 45%
- Comments- 30%
- Information from peers- 5%
- Personal experience-10%
- CCC subcommittee meetings-5%
- CCC whole group meetings-5%
- Other- 0%

Now, please fill in table below based on your experience:
APPENDIX A (continued)

<table>
<thead>
<tr>
<th>FACTOR</th>
<th>% CONTRIBUTION (total=100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milestones summary score</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>Information from peers</td>
<td></td>
</tr>
<tr>
<td>Personal experience</td>
<td></td>
</tr>
<tr>
<td>CCC subcommittee meetings</td>
<td></td>
</tr>
<tr>
<td>CCC whole group meetings</td>
<td></td>
</tr>
<tr>
<td>Other (please explain)</td>
<td></td>
</tr>
</tbody>
</table>

12) Having gone through the experience of assessing the residents’ performance on the milestones three times now, is there anything that:

   A) You, as an “evaluator” plan do differently for the next time (i.e. the next round of milestones evaluations)?

   B) You think the residency program should do differently for the next time?

13) Do you think the work of the CCC has helped identify struggling residents? Please explain.

14) How can the Department of Medicine help you in your role on the CCC?
APPENDIX B

CCC Focus Group-Interview Guide

Context:
- The interview will take place in the 1041 Kellogg (“small”) conference room in the Department of Internal Medicine on 1/20/16, 12:00-1:30pm.
- Media Services will set up the audio recorder to ensure high quality recordings.
- There are 2 parts to the session:
  12:00-12:30pm:
  Participants will complete a questionnaire while they eat lunch. The goals of this activity are to prime them for the focus group discussion, obtain another data source and collect some demographic data. They will be advised to do this individually. They will get a chance to share their responses during the interview.

  12:30-1:30pm:
  The focus group interview conducted by Dr. Elizabeth Baker.

Interviewer:
The interviewer should:
- Keep the session focused and well paced
- Draw out different perspectives
- Follow up on vague responses
- Try to facilitate a discussion amongst the participants as opposed to having them mostly interacting with you
- Maintain a respectful, comfortable and collegial environment during the interview
- Make sure that everyone has a chance to participate
- Explain the role of the observer
- Explain the purpose of the interview, which is to:
  - Give the participants an opportunity to discuss the topic and possibly reach new insights
  - Explore areas from previously collected data that require further explanation
- Explain the interview process i.e.
  - The semi-structured nature of the interview
  - Session will be audiotaped. Thus, participants should speak up and speak one at a time. Note: In the event of presentation or publication of the data, no identifiers will be used.
  - The session is purely voluntary
- Explain how the data will be used:
  - Information they provide will NOT be used to determine whether they can remain on the CCC.
  - The information obtained will ultimately be used to improve the quality of our assessment system.
APPENDIX B (continued)

Opening:
- Thank participants for being in the study and for agreeing to take part in the focus group.
- Explain the purpose of the interview and the research (see above).
- Confirm participation and remind participants that the interview will be audiotaped. They are under no obligation to participate.
- Explain the rules for the interview (see above).

Middle:
Ask the interview questions listed below-

1) “Having now gone through 4 reporting cycles, what do you feel is the role of the CCC?”
   - Primarily to identifying struggling residents?
   - Being able to identify all levels of performance?
   - Providing feedback to the program director
     - About struggling residents?
     - About all residents?
   - Be involved in improving the assessment system?
     - If so, how?
     - If not, please explain?
   - To see if residents are progressing along the milestones scale?
     - If yes, how should that occur? What should growth look like?
     - If no, please explain

2) “What is your perspective on the relative importance of the following factors in your decision making process?”
   - Comments (from evaluations)
   - Milestones summary score?
   - Your personal knowledge of a resident?
   - “Hearsay” (informal input from other faculty, residents or staff)?
   - CCC subcommittee discussion?
   - Large CCC meeting discussion?
   - Other?

3) Which milestones do you tend to have difficulty assessing? (Please refer to the milestones handout)? Why?
   - If the participants state that the problem is a lack of assessment data, probe to find out if that is indeed the issue or is it a lack of clear expectations for resident performance- or both?
   - Other factors? e.g. PC4 (procedural skills)

4) “Regarding the subcommittee structure of the CCC, what is your perspective on following a cohort of residents as they progress through their training versus staying focused on assessing residents at a particular level?”
APPENDIX B (continued)

5) “In an ideal world, what (if anything) would you want in order to do this job even more effectively?”
   - If more data, what data & why?
   - If more time, ask for an explanation (this may also shed light on what they think are the goals of the CCC)
   - If they want faculty development, what specifically do they want to learn more about?

6) “What do you think you have learned about the process of evaluating residents using the milestones framework that you did not know when we first started?”
   Is there anything that has become easier?
   Is there anything that has become harder?

7) Is there any feedback you would like to receive about your milestones scores?
   - If yes, please describe
   - If no, please explain

Closing:
   - Summarize the key points of the discussion to check accuracy
   - Give them a chance to add any other details to what they have already stated.
   - Thank them for their participation.
CITED LITERATURE

ACGME. ACGME Program Requirements for Graduate Medical Education in Internal Medicine

https://www.acgme.org/Portals/0/ACGMEClinicalCompetencyCommitteeGuidebook.pdf
Accessed 06/24/2016.


BIBLIOGRAPHY


Frambach JM, van der Vleuten CP, Durning SJ. AM last page. quality criteria in qualitative and quantitative research. Acad Med. 2013; 88:552.


NAME: Andem Ekpenyong, M.D.

Department of Internal Medicine
Rush University Medical Center
1653 West Congress Parkway
Chicago, IL. 60612
(312) 942-5269
(andem_ekpenyong@rush.edu)

EDUCATION:
National Diploma, Science Technology, The Polytechnic
Calabar, Nigeria, 1987-1989
B.S., Biology, Loyola University of Chicago 1989-1992
M.D., Rush Medical College, 1992-1996
Internal Medicine Residency, Rush Presbyterian St. Luke’s

TRAINING IN TEACHING:
Chief Resident, 1999-2000
Masters in Health Professions Education, University of
Illinois- Chicago, 2012-present
ACGME course titled “Developing Faculty Competencies
in Assessment,” March 2015
Certified in the Medical Education Research Certificate
(MERC), 2010
Stanford Faculty Development Program (SFDP)-Clinical
Teaching, 2007
AAMC’s Professional Development Seminar for Early
Career Women Faculty, 2005

ACADEMIC APPOINTMENTS:
Assistant Professor, 2001-2013
Associate Professor, 8/2013-present

EMPLOYMENT:
Rush University Internists, 6/00-present

CERTIFICATION:
Board Certified in Internal Medicine, 8/1999
ABIM Maintenance of Certification program, 5/2009

HONORS AND AWARDS:
Mark H. Lepper award (society of teachers), 2009
Teaching and Service, 1999-2000, 2002
Hero of the Heart, 2001
Best MHPE (Masters in Health Professions Education) Core Course Paper, 2014 Award (for Course 503 Curriculum & Evaluation)
Academic Mentoring Program (Department of Internal Medicine, RUMC), 2015-2016

SOCIETY MEMBERSHIPS: Society of General Internal Medicine (SGIM), 2005-present
Association of Program Directors in Internal Medicine (APDIM), 2007-present

TEACHING: M1 Introduction to the Patient Co-Course Director, 2002-2008
M2 Introduction to Clinical Skills, Co-Course Director, 2002-2008
Associate Program Director, 2007-present

COMMITTEES:
Departmental: Residency Curriculum Committee- Chair, 2000-present
University: Committee on Affirmative Action, 2003-2004

PROFESSIONAL MEMBERSHIP:
REGIONAL: Midwest Society of General Internal Medicine leadership board, 2011-2012
(At-Large member)
Midwest Society of General Internal Medicine- President Elect, 9/2012-9/2013
Midwest Society of General Internal Medicine-President, 9/2013-9/2014
University of Illinois, Department of Medical Education, 1/2015- present
Admissions Committee - member

University of Illinois, Department of Medical Education, 9/2015- present
Visibility Task Force - member

NATIONAL:

Association of Program Directors in Internal Medicine (APDIM), 2010-present
Program Planning Committee member, 2011-2014
APDIM nominating committee, 2013-2014
SGIM Board of Regional Leaders (BRL), 2013-2014

SGIM Program Planning Committee, 2014-2015

Alliance for Academic Internal Medicine (AAIM) Clinical Competency Committee Collaborative Learning Community (CCCCLC), Co-Chair, 1/2015-6/2016

AAIM Innovations Committee, 9/2015- present

ABSTRACTS:


Ekpenyong A, Uchida T. How Accurate are Medical Students in Self-Assessing Their Clinical Skills? JGIM 2007; 22 (suppl 1)

Ekpenyong A, Riddle JM, Uchida T, Baker E, Boyd K. How Do Third-Year Medical Students Identify Clinical Skills Deficiencies? JGIM 2009; 24 (suppl 1)
NON-PEER REVIEWED:


PRESENTATIONS:

Regional:
Oral Presentations:

Ekpenyong A, Uchida T. Students Perceptions of Preparedness for the Clinical Skills Assessment and Step 2 CS. Chicago Clinical Skills symposium at Northwestern University, 6/2006

Uchida T, Ekpenyong A. Medical Students Self-Assessment of Clinical Skills: revisited Chicago Clinical Skills symposium at Northwestern University, 7/2007


Uchida T. Ekpenyong A. Medical Student Self-Assessment of Clinical Skills. CGEA, 3/2007


Workshops:

Ekpenyong A, Uchida T, Boyd K, Baker E, Riddle J. What Do Third Year Medical Students Identify as Clinical Skills Deficiencies? 7th Annual Clinical Skills Education and Assessment Chicago Style Conference- at Loyola Univ. of Chicago Stritch School of Medicine. July 23, 2009 Also accepted for Midwest SGIM 9/09


Ekpenyong A, Uchida T, Riddle J, Heiman H, Boyd K, Baker E. Clinical Skills Remediation. CGEA, 4/9/10 and SGIM, Minneapolis, MN, 4/30/10


Invited Seminars and Lectures:


Ekpenyong A. Evaluating Clinical Faculty: Observation before Feedback. Invited by Dr. Susan Vanderberg-Dent, Designated Institutional Officer. Rush University Medical Center. Presented at the GME Leadership Development Program on 2/12/2013.

Ekpenyong A. Giving Feedback to Faculty. Invited by Dr. Susan Vanderberg-Dent, Designated Institutional Officer. Rush University Medical Center. Presented at the GME Leadership Development Program on 4/9/2013.
Invited by Dr. Joshua Baru (Program Director, Stroger/Rush Hospice and Palliative Fellowship) to deliver lecture on the “Learning Climate” to faculty in the divisions of hospital medicine and general medicine at Stroger Hospital on 10/8/14.

Invited Facilitator:

Invited by Dr. Cynthia Ledford (Associate Vice Chair of Medical Education, Department of Internal Medicine at Ohio State University College of Medicine) to participate as a trained facilitator in the delivery of the Stanford Faculty Development Clinical Teaching seminars. This content was delivered to 55 faculty across multiple disciplines, 6/10-6/11/14.

Invited by Dr. Eric Holmboe (Associate Vice President of Milestones Development, ACGME) to facilitate two 4-hour workshops on direct observation and feedback for faculty enrolled in the ACGME assessment course on 4/13/16 and 5/18/16.

National:

Posters:


Oral Presentations:


Workshops:


Rencic J, Chandler D, Ekpenyong A, Gentilesco B, Snydman L. Direct Observation of Resident Teaching During Work Rounds: Practical Tips to Implementing Direct Observation at your Institution. SGIM, Minneapolis, MN, 4/30/10

Ekpenyong A, Dedhia R, Gordon M, Litzelman D, Vannerson J. Direct Observation of Faculty: Faculty as Students. SGIM, Orlando, FL, 5/12/2012


Precourses/Mini-Courses:


(Organized this precourse)

Invited Seminars or Lectures:

(Presented as part of APDIM precourse. Invited by Drs. Cynthia Smith and Kelly Caverzagie), Las Vegas, NV. 4/2011


Invited by the APDIM Program Planning Committee to deliver a small group discussion on “Self Directed Learning” (along with Mark Shalaby, MD from Leigh Valley Health Network, PA) at the upcoming APDIM fall 2012 meeting

Ekpenyong, A. Direct Observation: Survey Data from Residents and Program Directors. Invited by the APDIM Program Planning Committee to be one of the plenary speakers at during the “Direct Observation” plenary at the fall 2013 APDIM meeting. New Orleans, LA 2013

International:
Posters:

Ekpenyong, A. Pilot Study: A Milestones-Based Self-Directed Learning (SDL) Survey for Internal Medicine Residents. AMEE, Prague, Czech Republic 2013 and ICRE, Calgary, Canada 2013


Ekpenyong A, Uchida T. How Accurate are Medical Students in Self-Assessing their Clinical Skills? SGIM, Toronto, Canada, 2007

Oral Presentations:

Ekpenyong A, Caverzagie K. Implementing the Internal Medicine Milestones in U.S. Residency Training. Accepted for oral presentation at The International
Conference on Residency Education in October 2012, Ottawa, Canada

Ekpenyong A. How Do Clinical Competency Committees Make Decisions about Internal Medicine Residents’ Achievement of the Milestones? A Pilot Study Using Grounded Theory Approaches. Also presented at International Conference on Residency Education (ICRE), 10/2015, Vancouver, CA.

REFERENCES:

Available upon request