Tower Health

Scholar Commons @ Tower Health

Reading Hospital Internal Medicine Residency

Internal Medicine Residency (Faculty and Residents)

12-15-2016

Design of a clinical competency committee to maximize formative feedback.

Anthony A. Donato

Richard Alweis

Suzanne Wenderoth

MEDICAL EDUCATION/MEDICAL STUDENT

Design of a clinical competency committee to maximize formative feedback

Anthony A. Donato, MD, MHPE[1]*, Richard Alweis, MD[2] and Suzanne Wenderoth, MD[1]

[1]Department of Medicine, Reading Health System, Reading, PA, USA; [2]Department of Medical Education, Rochester Regional Health System, Rochester, NY, USA

Background: As the next phase in the roll-out of Next Accreditation System, US residency programs are to develop Clinical Competency Committees (CCCs) to formally implement outcome-based medical education objectives in the resident assessment process. However, any changes to an assessment system must consider balancing formative and summative tensions, flexibility and standardization tensions, fairness and transparency to learners, and administrative burden for faculty.

Objectives/Methods: In this article, one program discusses the approach one internal medicine residency took to create a developmental model CCC. In this model, a learner's mentor presents the argument for competence to the CCC, while a second reviewer presents challenges to that argument to the rest of the committee members. The CCC members provide other insights and make recommendations. The mentor presents the final committee recommendations to that resident, who then works with the mentor to develop a plan for future action.

Results: CCC second reviewers spent an average of 30.4 min (SD: 11.4) preparing for each resident's discussion, a duty performed 5–7 times every 6 months. Faculty development was associated with an increase in the number of action-oriented comments in the meeting minutes (3.2–4.1 comments per resident, $p = 0.001$). CCC members and mentors gave higher Likert-type ratings than residents for fairness (4.8 vs. 4.0) and learning prioritization (4.7 vs. 4.2), but similar ratings for transparency (4.0 vs. 4.2).

Conclusion: Developmental model CCCs may be feasible for residency programs, but faculty development may be necessary.

Keywords: *internal medicine/education*; *education, graduate*; *medical/organization and administration*; *faculty*; *medical*; *clinical competence*

*Correspondence to: Anthony A. Donato, 6th and Spruce Streets, West Reading, PA 19610, USA, Email: Anthony.donato@readinghealth.org

Implementation of outcome-based medical education (OBME) through the roll-out of the Next Accreditation System (NAS) has brought forward a decade of rapid and seismic change for US residency training programs (1). Early steps of defining core competencies and delineation of dimensions of those competencies through the milestones project have now paved the way for the next critical iteration in OBME: incorporation of outcomes into competence determinations. In 2013, the Accreditation Council for Graduate Medical Education (ACGME) guided programs to create Clinical Competency Committees (CCCs) to fulfill this essential implementation step. This task requires all ACGME-accredited programs to delineate the purpose of their CCC and to develop processes for the fair and transparent determination of resident competence while aligning the CCC with their current assessment system (2). Individual programs were given considerable latitude to select and train members as well as to outline procedures for the aggregation and synthesis of resident data (2).

As assessment drives learning (3), the shift to OBME assessment paradigms is intended to guide training programs to produce members of an efficient, cost-effective team-based healthcare system (1). In addition, these changes were purported to offload the administrative burden of the prior accreditation system while possibly improving the efficiency of residency training through reduction of variability in educational processes (4). One counterargument to OBME, and more generally to all 'reductionist' models of assessment, is that any objectification in an assessment system risks trivialization of that system (5). Specifically, if either a learner or an assessor fails to see meaning in a parameter being assessed, the assessment activity may become a grading exercise

decoupled from its original learning objectives. Further, if learners see only the potential for negative summative judgment in formative feedback opportunities, feedback itself will become onerous to both parties and learning opportunities will be lost (6). Finally, some argue that the standardization of OBME may come at the cost of flexibility and individualization of training programs for specific learners (7). Therefore, the development of the goals and processes of a residency's CCC may have critical downstream implications to individual trainees and to the effectiveness of the educational program. In this article, we discuss the struggles and lessons learned from the creation of one CCC in a medium-sized internal medicine residency at an independent academic medical center.

The problem and ways to approach it

A 2015 qualitative analysis of 34 CCCs by Hauer et al. (8) found that most committees were oriented toward a 'problem-identification' model, in which the committee members' efforts focused on searching performance data for 'red-flags' to identify struggling residents. In this model, records of residents with performance issues were reviewed in detail for other events, while residents without issues were minimally reviewed, if at all. The members of problem-identification model committees noted concerns with the biases in decisions associated with reviewing data points that often relied on a few verbal or anonymous emails as sources. They also noted that residents were reluctant to receive the committee's formative feedback that they perceived to be summative, high-stakes judgments. However, a few programs pursued what was characterized as a 'developmental' model, in which CCCs compared all residents' performance data against established benchmarks, with focused discussion on areas for resident growth. Those using a developmental model identified concerns of increased time for the CCCs to apply the complex milestone format and the need for faculty development in qualitative methods as potential implementation challenges (8).

Potential pitfalls in CCC creation

Formative and summative tensions in assessment systems

Summatively directed assessment strategies and performance-oriented outcomes may have negative effects on learner motivation (9). Furthermore, grades themselves are poor carriers of feedback information and, in fact, inhibit any learning gains that formative feedback may provide (10). Standardization of assessment tools, especially if tools require assessors to translate human behaviors into numerical ratings, risks trivializing the assessment activity (6, 11). In contrast, narrative-based assessment systems require assessors to understand and apply qualitative research validity concepts including saturation, prolonged engagement, and triangulation (3). They also may require assessors to apply mixed methods to weigh disparate assessments (e.g., narrative-based direct observations and numerical in-training exam scores) against one another to create holistic synthetic judgments (4). Knowledge of the strengths and weaknesses of individual assessment tools as well as the rating idiosyncrasies of individual assessors themselves may be necessary for CCC members to separate 'signal' from 'noise' when making competence decisions.

Flexibility versus standardization tensions

By specifically describing behaviors that represent the goals of completed training, outcome-based paradigms may add standardization to an assessment process (7). However, any assessment standardization may inadvertently conceal important individual differences between trainees, making a program less flexible for its higher and lower performing learners (3). Furthermore, in more complex summative evaluation systems, formative feedback may be misperceived as summative which could threaten assessor–learner relationships, possibly resulting in range restriction or failure of faculty to record written constructive feedback (8).

Fairness and transparency tensions

Summative decisions on promotion can have long-lasting financial and career implications to learners. Decisions to place a learner on probation often come too late in training and can have serious and lasting negative effects on careers (12). Credible and trustworthy judgments must be made based on all available evidence and using transparent processes that can be audited by outsiders, if necessary. The learners themselves must feel that the process is credible, or the assessments generated, even if intended as formative and low stakes, will be trivialized by learners and assessors and learning gains will be lost (3).

Faculty administrative burden

Program directors note concerns with the time needed to aggregate and assess each resident on each of the 22 milestones (13, 14). Program directors now must teach other faculty members to apply new assessment processes as well as instruct their CCC members to aggregate these scores to create a synthetic judgment of resident competence (15). In addition, the number of observations required for defensible summative assessments of the milestones may create a significant burden on faculty (16). While improving information systems have been cited as a panacea to map assessment ratings to specific competencies, expert judgment is still needed to interpret the value and significance of each data point as it relates to others, considering context, rater, and level of training.

Our program

In 2013, Reading Health System formed a CCC with the purpose of guiding the growth of their 36 internal medicine residents. Residents are assessed primarily using a combination of scoreless, narrative-based single encounter assessments and summative evaluations with scored ratings (Table 1). The committee consists of six voting members, including the program director and two non-voting chief residents. Seven one-hour meetings are held every 6 months, with approximately 5.9 (range 2–8) residents discussed per meeting. All residents are discussed semi-annually, regardless of performance. The residents discussed in each meeting are those who are assigned to one specific institutional mentor. Mentors are assigned five to six mentees. That mentor meets three times a year with each of their mentees to review a mentee's progress by reviewing all assessments including all unlocked documents in their structured portfolio (20, 21). Not all mentors are members of the CCC, but if a mentor is a member of the CCC, that mentor attends the CCC as a non-voting advocate for that resident with respect to that individual mentee. During the meetings, mentors make the positive case for the competence of that resident. One other member of the committee (termed a 'second reviewer') also reviews all of the resident's performance data except the portfolio to identify any potential concerns in that resident's performance. The other CCC members listen to both arguments, then probe for more information from presenters or add their own observations. Voting on final determinations of competency occurs only in spring session and only in promotion decisions (graduating 1- and 3-year residents). Decisions are determined by majority, and the resident's mentor abstains from voting. Copies of the NAS milestones are contained in each resident folder and are used by reviewers as a benchmark, but the milestone ratings are determined later by the program director using the feedback of the CCC. Reliability and validity of feedback decisions are achieved by

Table 1. Evaluation tools used at Reading Health Internal Medicine Residency

Tool	Frequency (per resident)	Description
Reading Minicard Direct Observation Tool (17, 18)	40 per year	Prompted, scoreless narrative form with prompts for action plan based on direct observations of patient care activities
360-degree evaluation	8–15 per year	Single question, free-text narratives; performed by nurses, case managers, resident peers
Patient evaluations	3–5 per year	Seven questions, 3-point behavioral scale
End-of-month evaluations	12 per year	8–12 questions based on potentially observable milestones for that rotation, using same behavioral anchors
Rapid Response Simulation evaluation	5–12 (depending on mastery achievement) per 3 years, seniors only	3-domain, 8-question survey on behaviorally anchored 4-point scale performed by nurse observers; 2 questions with narrative from physician observer
Team meeting evaluation	2 per year	3-domain, 6-question survey by ambulatory team leader; behaviorally anchored 3-point scale with narrative summary
In-training exam	1 per year	300-question exam, with raw score and percentile rank by resident year
Subspecialty knowledge exams	4–6 per year	10–20 question internally developed pre- and post-tests of subject knowledge, generally MCQ format
OSCE evaluations	1 (10 station) per 3 years	Narrative direct observation performed using Minicard on 3 of 10, by residents and mentors
Evaluations of conference teaching (19)	4 per year	Three domains (planning, teaching techniques, presentation skills) rated using narrative descriptions
Quality improvement leadership direct observations	2 per year, seniors only	Prompted, three-domain, narrative-based form
Quality improvement reflections	2 per year	Completed by residents; 8 open-ended questions that demonstrate their understanding and contributions to group project
Reading Health e-portfolio (20)	Ongoing	Electronic repository of all above measures with resident reflections; also Curriculum Vitae (CV), Evidence-Based Medicine (EBM) searches, professional development plans; residents grant read-only access to mentors. Reviewed by mentor, not by CCC second reviewer

qualitative methods using triangulation, prolonged engagement, saturation, and member checking (with resident). Faculty development for the CCC is done in real time as part of the CCC meeting as well as in separate sessions, including 'Introduction to Milestones', 'Primer on Qualitative Methods', and 'Interpretation of Narratives' (a total of 3 h in the past 2 years). One facilitator keeps time, re-focuses discussions, and synthesizes the key recommendations. Those recommendations are then included in the minutes and delivered to the resident by the resident's mentor for the development of an action plan.

Program assessment

Outcomes were assessed by anonymous surveys of CCC participants and current residents, as well as qualitative analysis of the CCC minutes. The work was deemed to be Quality Improvement and exempt from Institutional Review Board (IRB). Qualitative analysis of the CCC was performed independently by two authors (AAD and RA). Unique utterances were characterized using a previously developed rating paradigm (17) as minimal (e.g., 'great job'), observational (e.g., 'nice connection with patients') or action oriented (e.g., 'work on handling family meetings'). Two investigators were trained in coding until 90% agreement was reached. Thirty percent of comments were double-coded to ensure reliability, with differences resolved by consensus. Reliability of double-coding was measured by modified kappa. Unique utterances were reported as an average number of comment types per resident during a single review. Reported rates were then compared in the 6-month periods before and after faculty development sessions. Comparisons between frequencies were performed using chi-square testing.

For the qualitative analysis of CCC minutes, 746 unique utterances were noted over the two time periods. Modified kappa agreement between reviewers was 93.0%. Results are presented in Table 2. Number of action-oriented comments increased significantly from year 1 to year 2 ($p = 0.001$).

CCC members, mentors, and residents were surveyed as to the fairness, transparency, and learner-centeredness of the assessment system. CCC members were also queried about their preparation time needed and also rated the usefulness of the assessment tools. Response rate was 100% from mentors and 69% from residents. Results of the survey are in Table 3. CCC participants noted that their second reviewer duties took 30.4 min (SD: 11.4) to prepare. CCC members noted that direct observations, in-training exam scores, and end-month rotation evaluations were most useful to inform their competence decisions, and 360-degree evaluations were least helpful.

Lessons learned

Our goal was to achieve a system which prioritized formative feedback without an undue administrative burden on faculty. We found that our system facilitates a closed loop of action-oriented feedback from the committee to the resident, and that higher quality (action-oriented) feedback has increased significantly following faculty development. Our findings show that a time-efficient developmental model CCC is feasible. However, our faculty had a deep understanding of the strengths and limitations of their local assessment tools and knew the idiosyncrasies of individual assessors, which may have assisted successful implementation. We believe that adding our mentors to the CCC in the role of resident advocates and shielding portfolios from CCC review protects the authenticity, while adding to the transparency and fairness of the CCC process. However, given that resident ratings for fairness were lower than the CCCs and were similar for transparency, more efforts may still be needed to clarify CCC processes to this resident cohort. It has also been proposed to broaden the membership of the CCC by adding faculty who can champion the process to residents. We chose to not directly rate the residents on the 22 milestones at the actual CCC meetings to preserve time efficiency. It is unknown whether errors or biases are introduced when program directors translate the narrative CCC minutes to milestone ratings.

Future directions

Medical education needs to be accountable to the public, but also to our trainees (22). If Graduate Medical Education (GME)'s goal is to preserve learner-centered, holistic approaches tailored to individual needs of learners ('assessment for learning') (5), we must focus efforts on training and facilitation of expertise of the assessors and

Table 2. Qualitative analysis of CCC minutes, by year*

Year (number of residents reviewed in time period)	Minimal N (% minimal per resident)	Observational N (% observational per resident)	Action oriented N (% action oriented per resident)
2013–2014 (n = 45)	21 (0.5)	295 (6.5)	94 (2.1)
2014–2015 (n = 40)	5 (0.1)	222 (5.5)	110 (2.8)

*Chi-square test of differences in frequencies between groups: $p = 0.001$.

Table 3. Anonymous survey results of participants, mentors, and residents regarding the fairness, transparency, and prioritization of learning of the assessment system

Question[a]	Mentors and CCC members	Residents
Please characterize your feelings on the *fairness* of the overall system of assessment of competence in this internal medicine program	4.80	4.04
Please characterize your feelings on the *transparency* of the overall system of assessment of competence in this internal medicine program	4.00	4.16
Please characterize how well the program *prioritizes your* (resident) *learning* in its assessment system[b]	4.70	4.20

[a]Scale: 1 = not at all (fair, transparent, learner-centered), 5 = very (fair, transparent, learner-centered); [b]resident survey question stem read 'prioritized your learning', CCC member survey read 'prioritized-resident learning'.

assessment systems, rather than individual assessment tools (3). Future directions of study of assessment systems should include metrics on the impact of assessment systems on learners' short-term learning gains following system feedback as well as on a learner's long-term self-directed learning, with iterative changes made to assessment systems to guide continuous improvement.

Acknowledgements

Conflict of interest and funding

References

1. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system – Rationale and benefits. N Engl J Med 2012; 366(11): 1051–6. doi: http://dx.doi.org/10.1056/NEJMsr1200117
2. CCC ACGME Faculty Development Guide. Available from: https://www.acgme.org/acgmeweb/Portals/0/ACGMEClinicalCompetencyCommitteeGuidebook.pdf [cited 21 September 2015].
3. van der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, et al. A model for programmatic assessment fit for purpose. Med Teach 2012; 34(3): 205–14. doi: http://dx.doi.org/10.3109/0142159X.2012.652239
4. Holmboe ES, Yaghmour N, Yamazaki K, Miller R, Edgar L, Conforti L, et al. Reflections on the first 2 years of milestone implementation. J Grad Med Educ 2015; 7(3): 506–11. doi: http://dx.doi.org/10.4300/JGME-07003043
5. Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach 2011; 33(6): 478–85. doi: http://dx.doi.org/10.3109/0142159X.2011.565828
6. van der Vleuten CPM, Schuwirth LWT, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: Building blocks for theory development. Best Pract Res Clin Obstet Gynaecol 2010; 24(6): 703–19. doi: http://dx.doi.org/10.1016/j.bpobgyn.2010.04.001
7. Hodges BD. A tea-steeping or i-Doc model for medical education? Acad Med 2010; 85(9 Suppl): S34–44. doi: http://dx.doi.org/10.1097/ACM.0b013e3181f12f32
8. Hauer KE, Chesluk B, Iobst W, Holmboe E, Baron RB, Boscardin CK, et al. Reviewing residents' competence: A qualitative study of the role of clinical competency committees in performance assessment. Acad Med 2015; 90(8): 1084–92. doi: http://dx.doi.org/10.1097/ACM.0000000000000736
9. Harlen W, Crick RD. Testing and motivation for learning. Assess Educ Princ Pol Pract 2003; 10(2): 169–207. doi: http://dx.doi.org/10.1080/0969594032000121270
10. Shute VJ. Focus on formative feedback. Rev Educ Res 2008; 78(1): 153–89. doi: http://dx.doi.org/10.3102/0034654307313795
11. Hodges B. The question of competence: Reconsidering medical education in the twenty-first century. 1st ed. Ithaca, NY: ILR Press; 2014.
12. Guerrasio J, Brooks E, Rumack CM, Christensen A, Aagaard EM. Association of characteristics, deficits, and outcomes of residents placed on probation at one institution, 2002–2012. Acad Med 2015; 91(3): 382–7. doi: http://dx.doi.org/10.1097/ACM.0000000000000879
13. Aagaard E, Kane GC, Conforti L, Hood S, Caverzagie KJ, Smith C, et al. Early feedback on the use of the internal medicine reporting milestones in assessment of resident performance. J Grad Med Educ 2013; 5(3): 433–8. doi: http://dx.doi.org/10.4300/JGME-D-13-00001.1
14. Swoboda TK. Graduate medical education reforms need to better serve program directors. Acad Med 2015; 90(9): 1182. doi: http://dx.doi.org/10.1097/ACM.0000000000000818
15. Pangaro LN. Two cheers for milestones. J Grad Med Educ 2015; 7(1): 4–6. doi: http://dx.doi.org/10.4300/JGME-D-14-00738.1
16. Norman G, Norcini J, Bordage G. Competency-based education: Milestones or millstones? J Grad Med Educ 2014; 6(1): 1–6. doi: http://dx.doi.org/10.4300/JGME-D-13-00445.1
17. Donato AA, Pangaro L, Smith C, Rencic J, Diaz Y, Mensinger J, et al. Evaluation of a novel assessment form for observing medical residents: A randomised, controlled trial. Med Educ 2008; 42(12): 1234–42. doi: http://dx.doi.org/10.1111/j.1365-2923.2008.03230.x
18. Donato AA, Park YS, George DL, Schwartz A, Yudkowsky R. Validity and feasibility of the minicard direct observation tool in one training program. J Grad Med Educ 2015; 7(2): 225–9.

19. Donato AA, Harris I. Use of portfolios for assessment of resident teaching skills. J Grad Med Educ 2013; 5(3): 476–80.

20. Donato AA, George DL. A blueprint for implementation of a structured portfolio in an internal medicine residency. Acad Med 2012; 87(2): 185–91. doi: http://dx.doi.org/10.1097/ACM.0b013e31823f40a2

21. Wald HS, Anthony D, Hutchinson TA, Liben S, Smilovitch M, Donato AA. Professional identity formation in medical education for humanistic, resilient physicians: Pedagogic strategies for bridging theory to practice. Acad Med 2015; 90(6): 753–60. doi: http://dx.doi.org/10.1097/ACM.0000000000000725

22. Weinstein DF. The elusive goal of accountability in graduate medical education. Acad Med 2015; 90(9): 1188–90. doi: http://dx.doi.org/10.1097/ACM.0000000000000834