ORIGINAL CONTRIBUTION

Making Sense of Trainee Performance: Entrustment Decision-Making in Internal Medicine Program Directors

Katherine A. Gielissen[a,*], Samantha L. Ahle[b], Thilan P. Wijesekera[a], Donna M. Windish[a], and Danya E. Keene[c]

[a]Department of Internal Medicine, Yale University School of Medicine, New Haven, CT; [b]Department of General Surgery, Yale University School of Medicine, New Haven, CT; [c]Yale University School of Public Health, New Haven, CT

Background: Competency-based assessment is an important but challenging aspect of residency education but determines trainees' progression towards the ultimate goal of graduation. Entrustment decision making has been proposed as a supplementary metric to assess trainee competence. This study explores the process by which Program Directors (PDs) make entrustment decisions in Internal Medicine (IM) training programs. **Study Design**: Purposive sampling was used to recruit PDs from ACGME-accredited IM training programs to participate in a semi-structured interview. We analyzed interviews using an iterative, grounded theory-based approach to allow identification of themes that define the process of trainee entrustment. **Results**: Sixteen PDs were interviewed. Qualitative analysis showed that PDs use a dynamic process to understand trainee entrustability and progression towards competence, including construction of assessment networks, comparing performance to expected trajectory of trainee competence development, and bidirectional filtering and weighing of assessment data. **Conclusions**: PDs serve as a central processor by which assessment data on trainees is filtered, weighted, and compared an expected trajectory, all to gain understanding of trainee performance. Assessment networks are crucial to understanding trainee competence. While expected trajectory is an important tool to determine how trainees are progressing, its continued use may inject bias into the assessment process and slow transition to true competency-based assessment.

BACKGROUND

Competency-based assessment (CBA) and the ultimate decision regarding readiness for independent practice remains a significant challenge for graduate medical education. Though promotion along the Accreditation Council for Graduate Medical Education (ACGME) milestones is the standard by which residents are promot-

*To whom all correspondence should be addressed: Katherine A. Gielissen, MD, MHS, 1450 Chapel Street, New Haven, CT 06511; Tel: 203-903-6091, Fax: 203-789-3007, Email: katherine.gielissen@yale.edu, ORCID iD: 0000-0001-7547-3118.

ed through training, resident assessment remains fraught with inconsistency and impacted by factors unrelated to trainee performance [1]. The process of progressive independence towards the ultimate goal of graduation has also met challenges as medical institutions strive to achieve transparency, safety, and accountability to the public [2,3]. Previous work shows that active supervision is important for patient safety and trainee education [4]. However, it is known that trainees struggle to progress in their independent decision making if provided too much oversight and supervision [5,6]. As such, concerns have arisen that increasing institutional oversight may produce trainees with little experience functioning independently [3,5,6].

In response to these challenges, entrustable professional activities (EPAs) have been proposed to operationalize CBA and provide a rationale and structure for graduated autonomy in medical training [7]. EPAs are clinical tasks that are independently executable, observable, and reflect one or more clinical competency. The revolutionary component of this assessment paradigm is the concept of "entrustment," the process of trusting a trainee to perform tasks in the clinical environment after they have demonstrated sufficient competence [8]. Previous work shows there are multiple factors that contribute to entrustment decisions in real clinical settings, including the supervisor's propensity to trust, trainee competence, clinical context, and the task itself [8].

Clinical supervisors do not necessarily engage in a systematic process when performing day to day assessment of trainees and often disagree on the types of activities trainees should be able to perform across stages of training [9,10]. Much of this heterogeneity is due to the complex social environment in which entrustment decisions take place, as well as diversity of trainee knowledge, attitudes, skills, and overall trustworthiness. However, there has been increasing understanding that supervisors themselves play an important part in entrustment decision-making (EDM), and that the amount of experience they have as educators affects the data used to determine trust, their approach to supervision, perspective on their role as supervisors, and confidence in their own clinical abilities [11,12].

In 2012, the Alliance for Academic Internal Medicine (AAIM) Educational Redesign Committee put forth a list of 16 end-of-training EPAs, which all internal medicine (IM) residents should be able to perform unsupervised prior to independent practice [13]. Although the EPAs elucidated by AAIM are not mandated for reporting purposes, they represent a significant opportunity to embed milestone-based assessment into the daily tasks expected of a fully competent internist. Despite these challenges, there has been much advancement in CBA through the creation of the Clinical Competency Committee (CCC), whose purpose is to create accountability toward the public at large. Through the composition and process of the CCC varies between and within institutions, it shares some core common goals: (1) developing a shared mental model of what trainee performance should "look like," (2) oversee the assessment tools within the program, (3) monitor residents' progress along a developmental trajectory, (4) identify residents "at risk" for not graduating, and (5) oversee programmatic composition and educational opportunities [14].

The purpose of this study was to explore perceptions and frameworks of trust, entrustability, and EDM paradigms used by Internal Medicine (IM) residency Program Directors (PDs). PDs occupy a unique role in residency training in that they are responsible for both ad hoc (*e.g.* formative, in the moment) and summative (*e.g.* readiness for graduation, licensing) entrustment decisions. The goal of this qualitative investigation was to explore the factors that influence the process by which clinical supervisors decide to trust a trainee along a continuum of developing competence with the ultimate goal of independent practice, and better understand how ad hoc and summative entrustment decisions are made in IM residency training. To date, no study has explored how PDs employ a framework to discern progression of trainees through a framework of entrustment.

METHODS

We conducted a multi-institutional qualitative study between March 2017 and November 2018. The Yale University Institutional Review Board deemed this study exempt. Using purposive sampling techniques, participants were invited via email from a list of potential candidates from our acquaintance network. These candidates were sent up to three requests for participation with personal outreach via phone or email if he or she did not respond. If this strategy failed, the PD was excluded from our study. During the interviews, participants were asked for contact information of other PDs who may be interested in participation in our study; these PDs were also invited to participate ("snowball technique").

A semi-structured interview was developed using a previously defined framework of entrustment decision making put forth by ten Cate *et al.* [7] The interview questions were pilot tested on eight PDs and Associate Program Directors (APDs) of both medical and surgical specialties at our institution. While feedback from and subsequent discussion of these pilot interviews informed further revision of our interview template, the data obtained was not included in our study sample. The final protocol asked the participants to describe their personal and institutional approach to entrustment decisions in IM trainees based on their experience(s) working individu-

Table 1. Participant characteristics.

Characteristic	n (%)
Male	11 (69%)
Years as a PD or APD	
< 10 years	5 (31%)
10-20 years	10 (62%)
> 20 years	1 (6%)
Personal Practice Setting	
Both inpatient and outpatient	11 (69%)
Inpatient only	4 (25%)
Outpatient only	1 (6%)
Program Size, total number of trainees	
Small, < 50	4 (25%)
Medium, 50-100	6 (38%)
Large, > 100	6 (38%)
Training Program Hospital Setting	
University	12 (75%)
Community Teaching	3 (19%)
Other	1 (6%)
Training Program Region[1]	
Mid Atlantic	6 (38%)
Midwest	2 (13%)
Mountain West	1 (6%)
New England	6 (38%)
Southern	1 (6%)

[1]As defined by the Society for General Internal Medicine (SGIM). PD = Program Director, APD = Associate Program Director.

ally with trainees and with the programmatic oversight of a PD.

After establishing the semi-structured interview questions, we added several quantitative questions to the beginning of our interview: 1) years as a PD or APD, 2) size of program, and 3) practice setting. The purpose of these questions was to better discern whether these factors altered the decision-making of program directors as our pilot interviews indicated that these factors influenced the entrustment process.

Interviews were conducted over the phone or in person by one of two trained interviewers from our study team (KAG, TW). After informed consent was obtained, interviews were audio recorded. Interview transcriptions were performed by Rev.com®; the accuracy of each transcript was verified by at least one member of the research team and all identifying information was removed prior to analysis.

Our study used an inductive approach to data analysis. As part of this procedure, three members of the research team (KAG, TW, SA) familiarized themselves with all transcripts by reading them several times. During this process, we identified preliminary entrustment frameworks, which were captured in brief analytic memos to facilitate discussion. Following discussion, each member of the team coded five interviews independently to generate an initial code book. These codes, their meanings and applications were further refined through an iterative process until final consensus on coding schema and their definitions was reached. Each interview was independently coded by at least two members of the research team; consistency of coding application between individuals was confirmed by KAG for all interviews. When discrepancies in coding dyads existed, these were resolved by a third coder.

RESULTS

We interviewed 16 IM PDs from a total of 37 invited PDs of ACGME-accredited internal medicine programs (41%). The mean length of interviews was 47 minutes (range 19-62 minutes). Theoretical saturation was reached after 11 interviews; all 16 interviews were included in data analysis.

Table 1 shows demographic data for study participants. Most participants were male (n = 10, 67%), had 10-20 years of experience as a PD or APD (n = 10, 67%), and practiced internal medicine in both inpatient and outpatient settings (n = 11, 73%). There was a relatively even distribution in small (< 50 trainees, n = 4, 27%), medium (50-100 trainees, n = 6, 40%), and large (> 100 trainees, n = 5, 33%) programs. Most participants led a Universi-

Table 2. Primary themes from qualitative interviews on entrustment decision making in IM PDs, their definitions and functions.

	Definition	Functions	Illustrative Quotes
Assessment Network	PDs develop working knowledge of the personnel and resources in their program to gain understanding of trainee performance. These moving parts form a web of competency and entrustability monitored and managed by the PD.	Early alert system for struggling trainees Triangulation of trainee performance	"We have a very good culture for early detection of people with problems or challenges. We pick those up quickly." "I may have heard amazing things consistently over a number of different evaluators across multiple different settings about one resident who I'm about to work with, and my suspicion that they're going to be fully trustable and allow them more autonomy early on, I have a much higher expectation that's going to play out."
Expected Trajectory	Residents progress along an expected curve of ability as they move through their training, such that supervisors have a sense of ability based on year of training or time of year ("contextual"). This can be seen through the lens established by past or current trainees ("comparative").	Understanding progress in the comparison to peers Understanding performance in context of medical practice	"I think probably the way a lot of these things go it is more of an either comparative, so, hey, other interns can do this and this person just hasn't seemed to grasp it." "I've had new interns who clearly are not where they should be even coming in as an intern, who are really having trouble feeling comfortable getting a history and making any sort of management decisions."
Interpretation of Assessment Data	PDs must sort, prioritize and interpret assessment data to better understand trainee performance. Part of this approach is considering the individual performing the assessment, where and when the assessment is performed, and variables that could affect the accuracy and validity of assessment data.	"Weighing" information points based on their utility and source Filtering out inputs that do not add to understanding of trainee performance	"And certainly when we have some of our most experienced people, our more core faculty, our preceptors who've worked with residents literally in thousands of encounters over time, if they're raising a trust concern…then we're going to weight that probably a bit more." "There are some people who tend to blow the whistle on multiple different residents over time… so when [they] raise an alarm, we pay attention to it and we try to triangulate and get more information from the other people who have worked with this person, but we also recognize this is somebody who raised a lot of alarms."

ty-associated training program (n = 11, 73%).

We found PDs used a dynamic process to understand trainee entrustability and progression towards competence. These processes included construction of assessment networks, comparing performance to expected trajectory of trainee competence development, and bidirectional filtering and weighing of assessment data. Descriptions of these themes and representative quotes are shown in *Table 2*. No significant differences in themes were noted between programs of different size, region, or amount of experience reported by the PD.

Construction of Assessment Networks

Program directors described a network as necessary to their work in assessment and evaluation of trainees. This group of individuals and resources was primarily comprised of parties providing both formal and informal assessment, such as faculty members, chief residents, and ancillary staff, but could also be informed through other inputs such as trainee peers. As one PD stated,

"It relies upon me knowing my faculty. And their own sort of nature of how they tend to report, evaluate."

It should be noted that while the above information sources were considered invaluable, PDs considered residency application materials, such as Medical Student Performance Evaluation (MSPE), letters of recommendation, or other materials provided through the Electronic Residency Application Service (ERAS) largely irrelevant or unhelpful in understanding the entrustability of incoming trainees. Interns were generally not presumed to have a baseline level of competence until they demonstrated knowledge, skills, and attitudes within the training program, regardless of their institution of origin. Aside from personnel resources (*e.g.* trusted faculty), the network involved working knowledge of the physical and cultural resources available at their institution, such as the various features (strengths and challenges) of rotations and practice settings. For example, PDs would take into account the "intensiveness" of a specific rotation when looking at assessment data.

Networks provide an opportunity to triangulate and corroborate information on individual trainees and serve as an information system feeding into the CCC and ultimately, the PD. However, despite efforts to construct networks, PDs also described significant barriers to obtaining constructive information on their trainees, including delays in written evaluations for progression/promotion, poor specificity in constructive feedback (more specifically, feedback on learners who were performing at levels below expected), and concerns that negative feedback from attendings would lead to significant personal or career repercussions on trainees. As such, PDs described

multiple "back channels" by which certain – often constructive – information was relayed through informal means (*e.g.* personal e-mail, "hallway conversations," phone calls). One PD described the channels as follows:

"So sometimes they are outside of the normal method of the electronic evaluation right? So you might get an email from a nurse or you might get a comment from a attending who worked with the resident over a weekend but they're not asked to formally evaluate the resident."

Of note, input from patients was largely absent from the entrustment network for the programs involved in this study.

PDs use assessment networks to leverage resources at their institution to ensure monitoring and ongoing assurance of trainee progress. Another function of the assessment network was the early detection and remediation of trainees who were struggling. Most PDs described a rich and multifaceted approach to struggling learners in order to provide them the best opportunities to "catch up" (see "Expected Trajectory" below), including special branches of the CCC to develop individualized learning plans, disseminate feedback to residents, and altering the training environment to meet learners needs (*e.g.* changing rotation structures or pairing struggling learners with particularly strong residents). At times, so many resources were invested in struggling learners, PDs described concern that residents in the "middle" (*e.g.* those who were performing adequately, but had mild-to-moderate deficiencies in a few specific domains) were not getting needed attention on specific, more difficult to detect areas of struggle.

Expected Trajectory of Trainee Competence

PDs acknowledged there existed a significant heterogeneity of ability within and between trainees at various levels of training. For example, most PDs agreed two residents at the same level of training at the same time of year could not necessarily be entrusted to do tasks at the same level of competence. Despite this widespread agreement, an "expected trajectory" of progression in competence was both implicitly and explicitly embedded in the language of how PDs described their trainees. This trajectory was often contextual (*e.g.* a resident who has not yet done an ICU rotation would not be entrusted to operate in the ICU independently) or comparative (*e.g.* implicit comparison between trainees at a given level of training or time of year) in nature. Expected trajectory took many forms in the language of the PDs we interviewed (*i.e.* "lagging behind," "he/she wasn't where I expected him/her at this level of training or time of year," "catching up") but was present in every interview to some degree, and often assisted decision-making in regard to

amount and degree of supervision provided for the trainee, even with no prior contact between trainee and supervisor. As one PD stated,

> "I think the biggest initial trust comes from what it says PGY-level on your badge."

In all interviews, trajectory was heavily based on training level, and this context would largely dictate much of trainee progression towards graduation. PDs indicated that this was often the default method by which progression was viewed:

> "It's mostly time-based in that you expect that interns after a year are going to progress into a more supervisory role as a second year."

Expected trajectory also assisted in milestone reporting by translating gestalt impressions to where residents "fell" on the milestone scale. Translation of performance to milestones resulted in a loss of nuance to understanding of trainee performance. This "quantification of ability" was accompanied by a loss of richness of information as the complexities of clinical performance were interpreted on a numerical or narratively weighted scale.

Filtering and Weighing Assessment Data

PDs, along with the CCC, must sort, prioritize, and interpret assessment data to better understand trainee performance. Part of this process includes consideration of the individual performing the assessment, where and when the assessment is performed, and variables that could affect the accuracy and validity of assessment data. As one participant stated,

> "Basically, sitting at the program level and looking over evaluations, talking to faculty across all the spectrum has vastly deepened my understanding of just how to weigh different individual data points."

As part of this approach, PDs recognize assessment patterns and trends within and between assessors, and when performance outliers are important to the interpretation of trainee performance. Oftentimes, intimate understanding of the individuals making up a PDs assessment network was key to interpretation of data inputs. For example, if a PD encountered a poor evaluation for an otherwise well-performing resident, he or she may take into account the faculty member who provided the assessment, the richness and specificity of the assessment itself, and the context of the situation prior to adding it to the overall summative assessment of that individual. Notably, "outlier" assessment data were not discounted outright: data with a significant concern about a trainee might be taken seriously, even if that information was not congruent with what was known about the trainee. Often this occurred when the information was provided by an individual well known in the assessment network.

All information inputs were reported as valuable, though most useful information came from firsthand experience with a trainee, followed by that from trusted faculty and chief residents. Information of prior performance also affected the PDs impression when working with a trainee personally for the first time. As one PD stated,

> "Based on my role as program director, I have a lot of information [from the assessment network] about all of our trainees before I ever work with them, so I have basically a preset probability that I can trust them going into it, and it's not to say that it's 100% for anybody, but I may have heard amazing things consistently over a number of different evaluators across multiple different settings about one resident who I'm about to work with, and my suspicion that they're going to be fully trustable and allow them more autonomy early on."

As part of this filtering and weighing process, PDs reported the CCC as an invaluable sounding board to further calibrate and interpret assessment data. This was particularly true for larger programs, where PDs did not always have an opportunity to work directly with all their trainees. PDs uniformly reported the CCC as an instrumental part of the assessment process as it provided structure as well as a forum to come to consensus on trainee performance.

DISCUSSION

This study identified multiple important factors by which entrustment decisions are made in IM residency training programs. The three major themes include the use of assessment networks to collect data on trainee performance, expected trajectories to put this performance in some context or framework, and the filtering and weighing process by which these data are interpreted. Our data builds on prior work describing the process by which attending physicians develop trust in their trainees and adds further understanding as to how summative entrustment decisions are made at the individual and programmatic level. Considering these findings, we envision the PD as a central repository by which assessment data is filtered, weighted, and compared to a trainee's expected trajectory, all to gain understanding of trainee performance.

Our data indicate IM PDs dedicate significant time and resources towards the construction of an assessment network, which allows them to better understand the performance of their trainees in their specific educational environment. Working knowledge of educational resources also facilitate methods to intervene on struggling learn-

ers. Importantly, our findings show the crucial role of the CCC in the overall process of assessment; PDs rely heavily on data on the network and use the CCC to better filter and weigh data on trainee competence. These findings correspond to prior qualitative studies, which show the CCC is a reliable forum to facilitate performance review [15,16]. However, despite these advantages, there remains significant issues with the quality of data presented to PDs, thereby necessitating "back channels" by which essential information makes its way to programmatic attention. These findings are corroborated with other studies, which showed PDs assign relative weight to certain data inputs on trainee performance [17]. Our study adds to this understanding by showing PDs have an active, working knowledge of their faculty when considering assessment data, and that glowing or concerning assessments may or may not necessarily be meaningful in this process. While it is difficult to extract the impact of such complex social interactions, further study is warranted to understand what factors are important in developing a working understanding of these assessment networks.

Some findings of this study contrast starkly with our qualitative exploration of entrustment decision-making in General Surgery PDs, where trainees were thought to have some baseline level of entrustment on entering residency training [18]. Our results suggest that unlike a surgical field, IM PDs do not presume some baseline competence on entrance into residency training. The reasons underlying the different approach to entering learners (*i.e.* interns) is unclear but may be due to the specific tasks entrusted to learners at the entry level of each field, or to a hidden curriculum enforced by supervising attending and senior learners. While avoiding this presumption of competence likely prompts earlier detection of struggling learners, it is possible that not understanding entering interns' skillsets could inject redundancy into residency training (*e.g.* confirming skillsets that a learner already has) or direct energy away from struggling learners while programs are attempting to understand the competence of all their new learners simultaneously. Efforts to address these issues are already underway; the American Association of Medical Colleges (AAMC) has proposed 13 Core EPAs for Entering Residency, which are thought to provide a more consistent framework to inform PDs of an entering resident's skillset [19]. Despite these efforts, and prior studies indicating the importance of educational handoffs during the transition from Undergraduate to Graduate Medical Training, robust educational handoffs are not widely employed [20].

Most compelling of our findings was the use of contextual and comparative proxies by which to determine trainee progression during their residency training. The language to describe this "expected trajectory" was very much embedded in time and stage of training, rather than how residents were progressing along their individual, time-independent courses. This finding contrast starkly with the goal of competency-based training, which proposes a framework by which learners progress through training in a manner that considers their specific professional development alone, rather than how long they have been in training [21]. The tendency to use expected trajectory represents a double-edged sword in the assessment space: it decreases cognitive load for assessors and provides a framework by which trainees can be considered for promotion based on the performance of their peers and predecessors but also inserts bias into the assessment process. This finding is most prominent when examining the relationship between training level and the amount of trust afforded to trainees, despite there being acknowledged heterogeneity of performance among trainees at similar levels. Further scrutiny into the prevalence and application of expected trajectory in residency training programs is necessary to understand how often these frameworks are applied.

There were several limitations to our study. First, while our population is a nationally representative sample encompassing multiple types of programs, it is unknown if it is representative of all themes present in the entrustment process for IM PDs. While thematic saturation was reached and additional interviews were performed and coded after saturation, a more extensive study involving PDs from other regions or institutions may yield additional themes important to the entrustment process. Lastly, as shown above the entrustment process involves many more stakeholders aside from PDs and likely has more factors relevant to understanding trainee competence; this study is only developed to understand the perspective of IM PDs.

IMPLICATIONS

This study expands on prior EDM work and adds to understanding of entrustment processes in IM residencies. While governing bodies such as the ACGME and AAIM recommend transition towards time-independent, competency-based assessment of trainees, our findings indicate that PDs use comparative and contextual factors and an expected trajectory of growth to determine where a trainee is in his or her progression. This phenomenon is perpetuated by ongoing challenges with obtaining accurate and useful data on trainee performance, particularly on trainees who are struggling. Continued work is needed to better perform more reliable assessments in the workplace.

REFERENCES

1. Williams RG, Dunnington GL, Mellinger JD, Klamen DL.

Placing constraints on the use of the ACGME milestones: a commentary on the limitations of global performance ratings. Acad Med. 2015 Apr;90(4):404–7.

2. Cianciolo AT, Kegg JA. Behavioral specification of the entrustment process. J Grad Med Educ. 2013 Mar;5(1):10–2.

3. Kennedy TJ, Regehr G, Baker GR, Lingard LA. Progressive independence in clinical training: a tradition worth defending? Acad Med. 2005 Oct;80(10 Suppl):S106–11.

4. Baldwin DC Jr, Daugherty SR, Ryan PM. How residents view their clinical supervision: a reanalysis of classic national survey data. J Grad Med Educ. 2010 Mar;2(1):37–45.

5. Piquette D, Tarshis J, Regehr G, Fowler RA, Pinto R, LeBlanc VR. Effects of clinical supervision on resident learning and patient care during simulated ICU scenarios. Crit Care Med. 2013 Dec;41(12):2705–11.

6. Babbott S. Commentary: watching closely at a distance: key tensions in supervising resident physicians. Acad Med. 2010 Sep;85(9):1399–400.

7. Ten Cate O, Hart D, Ankel F, Busari J, Englander R, Glasgow N, et al.; International Competency-Based Medical Education Collaborators. Entrustment Decision Making in Clinical Training. Acad Med. 2016 Feb;91(2):191–8.

8. Hauer KE, Ten Cate O, Boscardin C, Irby DM, Iobst W, O'Sullivan PS. Understanding trust as an essential element of trainee supervision and learning in the workplace. Adv Health Sci Educ Theory Pract. 2014 Aug;19(3):435–56.

9. Sterkenburg A, Barach P, Kalkman C, Gielen M, ten Cate O. When do supervising physicians decide to entrust residents with unsupervised tasks? Acad Med. 2010 Sep;85(9):1408–17.

10. Dijksterhuis MG, Voorhuis M, Teunissen PW, Schuwirth LW, ten Cate OT, Braat DD, et al. Assessment of competence and progressive independence in postgraduate clinical training. Med Educ. 2009 Dec;43(12):1156–65.

11. Sheu L, Kogan JR, Hauer KE. How supervisor experience influences trust, supervision, and trainee learning: a qualitative study. Acad Med. 2017 Sep;92(9):1320–7.

12. Kogan JR, Hess BJ, Conforti LN, Holmboe ES. What drives faculty ratings of residents' clinical skills? The impact of faculty's own clinical skills. Acad Med. 2010 Oct;85(10 Suppl):S25–8.

13. Caverzagie KJ, Cooney TG, Hemmer PA, Berkowitz L. The development of entrustable professional activities for internal medicine residency training: a report from the Education Redesign Committee of the Alliance for Academic Internal Medicine. Acad Med. 2015 Apr;90(4):479–84.

14. Andolsek K, Padmore J, Hauer KE, Holmboe E. Clinical competency committees. A guidebook for programs Chicago. The Accreditation Council for Graduate Medical Education; 2015.

15. 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 Aug;90(8):1084–92.

16. Hauer KE, Cate OT, Boscardin CK, Iobst W, Holmboe ES, Chesluk B, et al. Ensuring resident competence: a narrative review of the literature on group decision making to inform the work of clinical competency committees. J Grad Med Educ. 2016 May;8(2):156–64.

17. Ekpenyong A, Baker E, Harris I, Tekian A, Abrams R, Reddy S, et al. How do clinical competency committees use different sources of data to assess residents' performance on the internal medicine milestones?A mixed methods pilot study. Med Teach. 2017 Oct;39(10):1074–83.

18. Ahle SL, Gielissen K, Keene DE, Blasberg JD. Understanding Entrustment Decision-Making by Surgical Program Directors. J Surg Res. 2020 May;249:74–81.

19. Englander R, Flynn T, Call S, Carraccio C, Cleary L, Fulton TB, et al. Toward defining the foundation of the MD degree: core entrustable professional activities for entering residency. Acad Med. 2016 Oct;91(10):1352–8.

20. Wagner D, Lypson ML. Centralized assessment in graduate medical education: cents and sensibilities. J Grad Med Educ. 2009 Sep;1(1):21–7.

21. Iobst WF, Sherbino J, Cate OT, Richardson DL, Dath D, Swing SR, et al. Competency-based medical education in postgraduate medical education. Med Teach. 2010;32(8):651–6.