Home Field Advantage? Comparing the Quality of EPA Observations Completed On- vs Off-Service

Kaitlin Endres, MD
Timothy J. Wood, PhD
Nancy Dudek, MD, MEd
Warren J. Cheung, MD, MMEd

ABSTRACT

Background Increasing evidence suggests that supervisors exhibit different assessment behaviors for residents from their own discipline compared to those completing a rotation from another discipline. As programs of assessment rely on collecting robust performance data to inform high-stakes decisions about progress and promotion, it is important to examine the quality of such inputs.

Objective To compare the quality of workplace-based assessments (WBAs) of emergency medicine (EM) residents by EM and non-EM assessors.

Methods This retrospective database study compared the quality of WBAs using the Quality of Assessment of Learning (QuAL) score (range 0-5), a previously published measure of WBA quality that has demonstrated strong psychometric characteristics. Five entrustable professional activities (EPAs), 3 procedural and 2 non-procedural, mapped to both EM and non-EM rotations, were selected for inclusion. Two hundred and fifty WBAs (50 WBAs per EPA; 25 EM and 25 non-EM), completed from July 2019 to June 2021, were rated by 3 blinded EM physician raters. QuAL scores were analysed using factorial ANOVA.

Results Mean QuAL scores for WBAs completed during EM rotations were significantly higher compared to those completed during non-EM rotations (3.66±0.99 vs 3.02±0.99). Further, mean QuAL score for procedural EPAs was significantly higher than non-procedural EPAs (3.61±1.00 vs 3.16±1.03).

Conclusions In this study, the quality of WBAs completed for EM residents during non-EM rotations was of lower quality compared to assessments on EM rotations.

Introduction

Context-rich narratives of performance and ratings that differentiate between a resident's strengths and areas for improvement are important springboards for actionable feedback for learning, and they contribute meaningful data to inform competence committee decisions about resident progress.[1-4] However, the validity and defensibility of such decisions are only as good as the performance data acquired to inform them.[5-9] There is increasing literature to suggest that supervisors observe and assess residents from their own discipline (on-service) differently compared to rotating residents from other disciplines (off-service).[10,11] Perhaps this relates to a supervisor's tacit sense of educational responsibility for residents within their own discipline, or maybe it relates to Telio's conceptualization of the "educational alliance"[12,13] where feedback is perceived as higher quality when the learning goals of supervisor and resident are aligned. Regardless, the majority of literature describing supervisor assessment behaviors for on- versus off-service residents focuses on longitudinal In-Training Evaluation Reports that take place over weeks or months of supervision rather than from individual patient encounters such as those reflected in entrustable professional activities (EPAs).[5-9] Additionally, the impact of supervisors' assessment behaviors on the quality of assessments generated for on- and off-service residents has not been formally evaluated.[10]

Junior residents, postgraduate year (PGY) 1 through 3, in emergency medicine (EM) may spend up to 40 to 50% of their training completing off-service rotations in other disciplines.[14] Off-service rotations may vary in duration but are typically 4 weeks in length, with the intention of providing increased exposure to specific patient populations and conditions, while offering clinical experiences intended to highlight the breadth and scope of practice of a particular discipline. EM residents complete the vast majority of their off-service rotations during PGY-1-3 of their training program. Given the considerable time residents spend off-service and the amount of performance data generated during these rotations, an examination of the

DOI: http://dx.doi.org/10.4300/JGME-D-24-00719.1

Editor's Note: The online supplementary data contains and example of the entrustable professional activities (EPA) assessment in context, the Quality of Assessment of Learning Score, and a list of selected EPA assessment descriptions and applicable off-service disciplines.

differences in the quality of EPA assessments collected on-service versus off-service is warranted—particularly in light of the literature suggesting that supervisors exhibit different assessment behaviors when supervising off-service residents.[10,15] Therefore, the objective of this study was to compare the quality of assessments completed for residents while on-service versus off-service.

KEY POINTS

What Is Known?
Supervisors may assess residents differently, depending on whether the resident is from their own discipline or another. It would be helpful to know if the quality of feedback differs in these situations, which is a gap in the current literature.

What Is New?
Using the Quality of Assessment for Learning (QuAL) score, this study of assessments of emergency medicine (EM) residents found that the quality of workplace-based assessments (WBAs) from EM rotations was higher than those from non-EM rotations.

Bottom Line
Program directors should be cautious in relying equally on WBA data from off-service rotations when making high-stakes decisions, and may need to employ calibration efforts for faculty across disciplines.

Methods

Study Setting and Participants

This retrospective database study compared the quality of on-service and off-service EPA assessments completed for EM residents as measured by the Quality of Assessment of Learning (QuAL) score.[16] This study was conducted in the Department of Emergency Medicine at the University of Ottawa, which is comprised of 2 tertiary care campuses of The Ottawa Hospital and includes approximately 50 Royal College of Physicians and Surgeons of Canada (RCPSC) EM residents. Included EPA assessments were completed by clinical supervisors for residents between July 2019 and June 2021 using a standardized EPA assessment form (see online supplementary data appendix A).

Measure of Assessment Quality

The formerly published QuAL score was designed using Messick's validity framework to measure the quality of short workplace-based assessments (WBAs), such as EPA assessments, that include narrative comments and a single global performance score.[16] See online supplementary data appendix B for a detailed breakdown of the QuAL score components. The QuAL score has demonstrated strong psychometric characteristics, including reliable scores and the ability to discriminate the quality of EPA assessments based on utility.[16-18] The QuAL score has previously been used to study the impact of supervisor-trainee continuity on the quality of workplace-based assessments as well as the influence of direct versus indirect observation on assessment quality.[17,18] Furthermore, the QuAL score has demonstrated high positive correlations with perceptions of assessment utility by residents, academic advisors, and competence committee members, further adding to its validity evidence.[19] The maximum QuAL score is 5, indicating high quality, and the minimum is 0, indicating low quality.[16] Based on the original QuAL derivation study, a 0.5 difference in QuAL score is considered educationally meaningful.[16]

Outcomes Measured

Multiple EPA assessments within the RCPSC EM curricular design are completed for residents during their time both on- and off-service. These WBAs are completed by the resident's clinical supervisor for that rotation. We selected 5 EPAs that represent a breadth of procedural and non-procedural tasks that are commonly assessed for residents both on- and off-service.[14,19] To be included in the study, at least 10% of the assessments for a particular EPA must have been completed during off-service rotations. We felt that if an EPA was completed off-service less than 10% of the time, this was insufficient for proper comparison in our study. Selected EPAs included clinical tasks that can be observed and assessed on a variety of off-service clinical rotations, including anesthesiology, cardiology, critical care, general internal medicine, general surgery, orthopedic surgery, neurology, pediatric emergency medicine, pediatric critical care, psychiatry, and trauma. See online supplementary data appendix C for a list of selected EPAs, their descriptions, and how they map to various off-service rotations.

Assessment forms for each of the 5 EPAs completed during the study period for RCPSC EM residents were downloaded from our electronic data management system by a research assistant and grouped into "on-service" and "off-service." Based on the included sample size calculation below, 25 on-service WBAs and 25 off-service WBAs were randomly selected for each of the 5 EPAs using a random number generator for a total of 250 forms.[20] Each completed WBA form was de-identified with regard to the rotation on which it was completed, the resident being assessed, and the supervisor completing the assessment. Each form was assigned a unique study number. Three blinded physician raters independently scored the quality of each completed WBA form using the QuAL score. Use of the QuAL score does not require raters to undertake any training other than to read the instructions on the form prior to completion of ratings. Based on previous studies[16-18] an acceptable reliability (>0.8) can be achieved with 3 raters.

Sample Size

For this study, the main comparison was differences in the mean QuAL score between on-service and off-service EPA assessments. Based on previous studies,[16-18] to detect a significant difference with a moderate effect size of 0.80, assuming a level of significance of P=.05, power of 0.80, and a standard deviation of 0.95, 50 EPA assessments (25 on-service and 25 off-service) would need to be rated for each EPA.

Analysis of Outcomes

Quality of EPA Assessment by On- Versus Off-Service: A comparison of mean QuAL scores between the on-service and off-service groups was conducted. QuAL scores were analysed using a factorial analysis of variance (ANOVA) with mean QuAL score as the dependent variable, and type of observation (on- versus off-service) and EPA type (1-5) as the independent variables. Further, QuAL scores for procedural versus non-procedural EPAs were compared using a between-subject ANOVA with mean QuAL score again as the dependent variable and procedural versus non-procedural EPA as the independent variable.

Reliability: The reliability of the QuAL score for EPA assessments was calculated using a generalizability analysis. In this model, individual EPA assessment forms were considered the object of measurement with condition (on- versus off-service) and EPA type (1-5) treated as between-subject factors. EPA assessment forms were nested within the EPA type and condition and crossed with physician rater (1-3). Mean QuAL score was the dependent measure. The variance components that resulted from these analyses were used to determine the reliability of the QuAL instrument.

This study has received full ethics review exemption from the Ottawa Health Science Network Research Ethics Board.

Results

A total of 250 EPA assessments were scored using the QuAL score by 3 independent staff physician raters. The mean QuAL score for EPAs assessed on-service was significantly higher than off-service (3.66±0.99 vs 3.02±0.99, P<.001; F(1240)=27.78, P<.001, η_p^2=.10) as demonstrated in FIGURE 1.

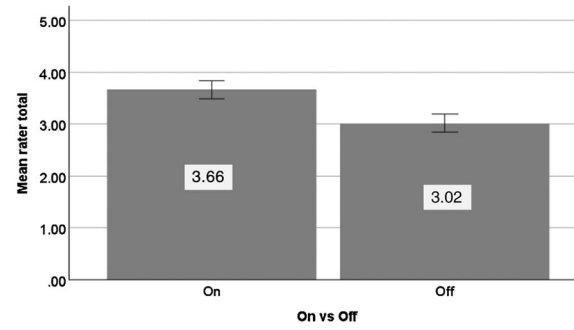

FIGURE 1
Mean QuAL Score for On-Service Versus Off-Service EPA Assessments

Abbreviations: QuAL, Quality of Assessment of Learning; EPA, entrustable professional activity.

Post-hoc analysis, which can be visualized in the TABLE, demonstrated that on-service QuAL scores were higher than off-service scores for each EPA. This difference was statistically significant for 3 of 5 EPAs: managing critical illness (C1), airway management, (C3) and managing psychiatric emergencies (C9).

As demonstrated in FIGURE 2, the mean QuAL score for procedural EPAs (C3, C13) was significantly higher than the mean QuAL score for non-procedural EPAs (3.61±1.00 vs 3.16±1.03, P<.001, F(1248)=11.58, d=0.44). This was the case regardless of whether EPAs were evaluated on- or off-service. The reliability of the ratings on the QuAL with 3 raters was 0.79 and 0.55 with a single rater.[21]

Discussion

Workplace-based assessments collected on- and off-service support individual learner growth and contribute meaningful performance data to inform high-stakes decisions made by Clinical Competency Committees (CCC). However, the validity and defensibility of these progress decisions are only as good as

TABLE
Summary of Mean On-Service vs Mean Off-Service QuAL Score for 5 Studied EPAs (Core 1, 3, 5, 9, 12)

EPA	Mean On-Service QuAL	SD	95% CI	Mean Off-Service QuAL	SD	95% CI	*P* value
Core 1: Critical Illness	3.57	1.07	3.13-4.01	2.67	1.01	2.25-3.09	.003[a]
Core 3: Airway Management	4.09	0.9	3.72-4.46	3.13	0.94	2.75-3.52	<.001[a]
Core 5: Emergent Medical and Surgical Management	3.39	0.92	3.01-3.76	3.11	1.04	2.68-3.54	.32
Core 9: Psychiatric Emergencies	3.48	0.97	3.08-3.88	2.75	0.88	2.83-3.39	.007[a]
Core 13: Advanced Procedures	3.77	0.99	3.36-4.18	3.42	0.94	3.04-3.82	.21

[a] Identifies statistically significant results.
Abbreviations: QuAL, Quality of Assessment of Learning; EPA, entrustable professional activity.

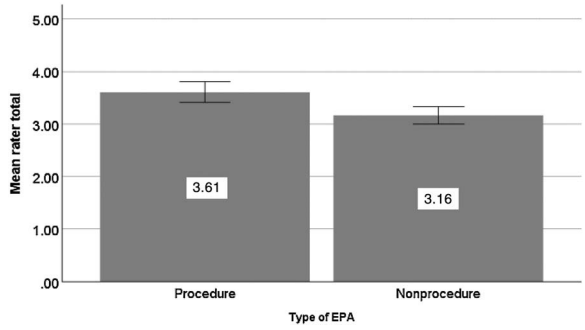

FIGURE 2
Mean QuAL Score for Procedural Versus Non-Procedural EPAs

Abbreviations: QuAL, Quality of Assessment of Learning; EPA, entrustable professional activity.

the quality of the inputs. Recent work has suggested that supervisors exhibit different assessment behaviors when assessing on-service versus off-service residents, leading to questions about the quality of off-service workplace-based assessments.

Our study found that the quality of on-service EPA assessments was significantly higher than off-service EPA assessments, as measured by the QuAL score. The observed difference has important educational implications, as lower quality assessments may be perceived by residents as less effective in driving learning and may not provide sufficiently detailed or accurate assessments of their performance to inform high-stakes decisions about progress and promotion made by the CCC.[19] Telio and colleagues' concept of the educational alliance[12,13] may offer some insights into these findings. The educational alliance was conceptualized as a framework for understanding what strengthens and weakens the supervisor-trainee relationship by drawing parallels to the therapeutic alliance or physician-patient relationship within psychotherapy.[12] The educational alliance notes that feedback is perceived as higher quality and therefore more valuable to the trainee when the supervisor is more engaged as an educator, committed to promoting growth of the trainee and when the learning goals of the trainee and supervisor are aligned.[13] In other words, feedback is perceived as higher quality by the trainee when the supervisor is seen as a partner in the educational alliance.[13]

Supervisors likely have a tacit sense of educational responsibility for residents within their own training program.[10] Additionally, having gone through similar training, supervisors are better able to understand and relate to the educational outcomes and training challenges of a resident within their discipline, which facilitates the provision of meaningful and actionable written feedback for improvement.[10,12,13] On the contrary, supervisors working with a resident outside of their discipline may find it challenging to make judgments of performance, particularly when they are unclear of the standard of competence expected of off-service residents.[3,22-26] Additionally, the learning goals of an off-service resident and the rotation-specific knowledge and skills relevant to their eventual practice may not always be evident to the supervisor. This lack of educational alliance can make it very challenging to offer contextually relevant feedback to and assessments of the trainee.[12,13,27,28] Thus, our findings highlight the importance of ensuring that each clinical experience (including off-service rotations) has clear learning outcomes, that assessments are designed specifically for these outcomes, and that these outcomes are clearly communicated to those who will be assessing the trainee.

Our study also found that the quality of assessments for procedural EPAs was significantly higher than for non-procedural EPAs regardless of the clinical setting. This is supported by previous literature that suggests that it may be easier for supervisors to provide specific and actionable feedback about observed procedural skills.[29-31] Procedural skills are more likely to be observed directly, leading to less reliance on surrogate data to draw inferences, promoting more authentic judgements of competence and facilitating increased trust between supervisor and trainee.[17,18,32-34] Further, procedural skills are traditionally assessed using checklists, which can facilitate the provision of specific and actionable feedback focused on the areas for improvement.[35,36] Lockyer and colleagues argued that assessment strategies focusing on "reliable checklists" help improve the quality of feedback and should be used even outside of observed structured clinical examinations.[37]

Limitations

This study has some important limitations to consider. Although EPA assessment forms were deidentified with regard to the clinical context of the EPA assessment, there may have been subtle cues in the feedback language that might have implied whether the assessment was completed on- vs off-service. Further, this study was conducted at a single center in a single RCPSC EM program, and the results may have been influenced by local cultural norms and assessment patterns. Data was collected during the COVID-19 pandemic, with social distancing and increased clinical workload which may have impacted overall assessment quality. The finding that off-service assessments were of lower quality than on-service assessments, however, has previously been observed in a pre-pandemic study, and thus is not likely a COVID-19 specific phenomenon.[10] Further, we did not study whether trainee gender or seniority affected the quality of feedback

received, and we did not evaluate the difference between on- and off-service EPA assessment quality for all EPAs in our national assessment framework. Instead, we selected a sample of 5 EPAs representing procedural and non-procedural tasks. Results may have differed with a different sample of EPAs. Finally, discipline-specific assessment quality was not evaluated in this study.

Future Directions

Our findings represent an educational gap and point to an opportunity to improve assessment practices. Future research should explore the perspectives of supervisors to better understand the challenges they face when providing feedback and assessment of off-service trainees. We offer that qualitative methods are needed to explore in greater depth the trainee-supervisor relationship and how the educational alliance may impact supervisors' assessment behaviors. Additionally, we aim to compare how the quality of feedback provided for the same resident differs on-service compared to off-service. These findings may inform faculty development initiatives that aim to better orient faculty to the specific educational needs of off-service trainees as well as the assessment needs of the sending training program. Finally, greater attention should be directed at determining ways of optimizing resident assessment during off-service rotations by designing assessment tools in a manner that aligns with the expectations and goals of the off-service rotation.

Conclusions

In our study, the quality of workplace-based assessments completed during off-service rotations was of lower quality compared to assessments completed during on-service rotations.

References

1. Royal College of Physicians and Surgeons of Canada. CanMeds 2015 Physician Competency Framework. Accessed January 23, 2024. https://canmeds.royalcollege.ca/en/framework
2. Watling CJ, Ginsburg S. Assessment, feedback and the alchemy of learning. *Med Educ*. 2019;53(1):76-85. doi:10.1111/medu.13645
3. Lefroy J, Watling C, Teunissen PW, Brand P. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. *Perspect Med Educ*. 2015;4(6):284-299. doi:10.1007/s40037-015-0231-7
4. Alquraini M, Baig L, Magzoub M, Omair A. Evaluation of off-service rotations at National Guard Health Affairs: results from a perception survey of off-service residents. *J Fam Community Med*. 2013;20(2):123-129. doi:10.4103/2230-8229.114773
5. Dudek N, Gofton W, Rekman J. Faculty and resident perspectives on using entrustment anchors for workplace-based assessment. *J Grad Med Educ*. 2019;11(3):287-294. doi:10.4300/JGME-D-18-01003.1
6. Rekman J, Gofton W, Dudek N, Gofton T, Hamstra SJ. Entrustability scales: outlining their usefulness for competency-based clinical assessment. *Acad Med*. 2016;91(2):186-190. doi:10.1097/ACM.0000000000001045
7. Dolan BM, O'Brien CL, Green MM. Including entrustment language in an assessment form may improve constructive feedback for student clinical skills. *Med Sci Educ*. 2017;27(3):461-464. doi:10.1007/s40670-017-0434-0
8. Cohen GS, Blumberg P, Ryan NC, Sullivan PL. Do final grades reflect written qualitative evaluations of student performance? *Teach Learn Med*. 1993;5(1):10-15. doi:10.1080/10401339309539580
9. Speer AJ, Solomon DJ, Ainsworth MA. An innovative evaluation method in an internal medicine clerkship. *Acad Med*. 1996;71(1):76-78. doi:10.1097/00001888-199601000-00048
10. Cheung WJ, Dudek NL, Wood TJ, Frank JR. Supervisor-trainee continuity and the quality of work-based assessments. *Med Educ*. 2017;51(12):1260-1268. doi:10.1111/medu.13415
11. Rakowski RT, Wightman JM, Hamilton GC. Objectives to direct the training of emergency medicine residents on off-service rotations. *J Emerg Med*. 1995;13(6):823-829. doi:10.1016/0736-4679(95)02032-2
12. Telio S, Ajjawi R, Regehr G. The 'educational alliance' as a framework for reconceptualizing feedback in medical education. *Acad Med*. 2015;90(5):609-614. doi:10.1097/ACM.0000000000000560
13. Telio S, Regehr G, Ajjawi R. Feedback and the educational alliance: examining credibility judgements and their consequences. *Med Educ*. 2016;50(9):933-942. doi:10.1111/medu.13063
14. Sherbino J, Bandiera G, Doyle K, et al. The competency-based medical education evolution of Canadian emergency medicine specialist training. *Can J Emerg Med*. 2019;22(1):95-102. doi:10.1017/cem.2019.41
15. Shofler D, He A, Lin TL, Chuang CT. An evaluation of off-service rotations in podiatric medicine and surgery residency training. *J Foot Ankle Surg*. 2019;58(3):480-483. doi:10.1053/j.jfas.2018.09.026
16. Chan TM, Sebok-Syer SS, Sampson C, Monteiro S. The Quality of Assessment of Learning (QuAL) Score: validity evidence for a scoring system aimed at rating short, workplace-based comments on trainee performance. *Teach Learn Med*. 2020;32(3):319-329. doi:10.1080/10401334.2019.1708365
17. Singh S, Cheung WJ, Dewhirst S, Wood TJ, Landreville JM. The influence of clinical coaching teams on quality of entrustable professional activity assessments. *AEM Educ Train*. 2023;7(3):e10879. doi:10.1002/aet2.10879

18. Landreville JM, Wood TJ, Frank JR, Cheung WJ. Does direct observation influence the quality of workplace-based assessment documentation? *AEM Educ Train*. 2022;6(4):e10781. doi:10.1002/aet2.10781
19. Woods R, Singh S, Thoma B, et al. Validity evidence for the Quality of Assessment for Learning score: a quality metric for supervisor comments in competency based medical education. *Can Med Educ J*. 2022;13(6):19-35. doi:10.36834/cmej.74860
20. Stat Trek. Random number generator. Accessed Janurary 15, 2022. https://stattrek.com/statistics/random-number-generator.aspx
21. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30-46. doi:10.1037/1082-989X.1.1.30
22. Watling C, Driessen E, van der Vleuten CPM, Vanstone M, Lingard L. Beyond individualism: professional culture and its influence on feedback. *Med Educ*. 2013;47(6):585-594. doi:10.1111/medu.12150
23. Yarris LM, Linden JA, Gene Hern H, et al. Attending and resident satisfaction with feedback in the emergency department. *Acad Emerg Med*. 2009;16(suppl 2):76-81. doi:10.1111/j.1553-2712.2009.00592.x
24. Jensen AR, Wright AS, Kim S, Horvath KD, Calhoun KE. Educational feedback in the operating room: a gap between resident and faculty perceptions. *Am J Surg*. 2012;204(2):248-255. doi:10.1016/j.amjsurg.2011.08.019
25. Balmer DF, Giardino AP, Richards BF. The dance between attending physicians and senior residents as teachers and supervisors. *Pediatrics*. 2012;129(5):910-915. doi:10.1542/peds.2011-2674
26. Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales. *Med Educ*. 2011;(45):560-569. doi:10.1111/j.1365-2923.2010.03913.x
27. Williams RG, Klamen DA, McGaghie WC. Cognitive, social and environmental sources of bias in clinical performance ratings. *Teach Learn Med*. 2003;15(4):270-292. doi:10.1207/S15328015TLM1504_11
28. Macrae CN, Bodenhausen GV. Social cognition: thinking categorically about others. *Annu Rev Psychol*. 2000;51:93-120. doi:10.1146/annurev.psych.51.1.93
29. Cheung WJ, Patey AM, Frank JR, Mackay M, Boet S. Barriers and enablers to direct observation of trainees' clinical performance: a qualitative study using the theoretical domains framework. *Acad Med*. 2019;94(1):101-114. doi:10.1097/ACM.0000000000002396
30. Kogan JR, Conforti LN, Bernabeo EC, Durning SJ, Hauer KE, Holmboe ES. Faculty staff perceptions of feedback to residents after direct observation of clinical skills. *Med Educ*. 2012;46(2):201-215. doi:10.1111/j.1365-2923.2011.04137.x
31. LaDonna KA, Hatala R, Lingard L, Voyer S, Watling C. Staging a performance: learners' perceptions about direct observation during residency. *Med Educ*. 2017;51(5):498-510. doi:10.1111/medu.13232
32. Hauer KE, Holmboe ES, Kogan JR. Twelve tips for implementing tools for direct observation of medical trainees' clinical skills during patient encounters. *Med Teach*. 2011;33(1):27-33. doi:10.3109/0142159X.2010.507710
33. Shoenberger J, Taira T, Tabatabai R, Osterman J. Protecting faculty time for direct observation shifts in a large emergency medicine residency program. *West J Emerg Med*. 2015;16(suppl):28-29.
34. Urquhart LM, Rees CE, Ker JS. Making sense of feedback experiences: a multi-school study of medical students' narratives. *Med Educ*. 2014;48(2):189-203. doi:10.1111/medu.12304
35. Regehr G, MacRae H, Reznick RKS, D. Comparing psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med*. 1998;73(9):993-997. doi:10.1097/00001888-199809000-00020
36. Ross V, Mauksch L, Huntington J, Beard JM. Interdisciplinary direct observation: impact on precepting, residents, and faculty. *Fam Med*. 2012;44(5):318-324.
37. Lockyer J, Carraccio C, Chan MK, et al. Core principles of assessment in competency-based medical education. *Med Teach*. 2017;39(6):609-616. doi:10.1080/0142159X.2017.1315082

Kaitlin Endres, MD, is a Senior Emergency Medicine Resident, Department of Emergency Medicine, University of Ottawa, Ottawa, Ontario, Canada; **Timothy J. Wood, PhD,** is a Full Professor, Department of Innovation in Medical Education, University of Ottawa, Ottawa, Ontario, Canada; **Nancy Dudek, MD, MEd,** is a Full Professor, Division of Physical Medicine & Rehabilitation, Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada; and **Warren J. Cheung, MD, MMEd,** is an Associate Professor, Department of Emergency Medicine, University of Ottawa, Ottawa, Ontario, Canada.

Funding: This study was funded by a grant from the Canadian Association of Emergency Physicians, Physician Services Incorporated, and the University of Ottawa Centre for Innovation in Medical Education.

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

This work was previously presented at the Society for Academic Emergency Medicine Annual Meeting, May 14-17, 2024, Phoenix, Arizona, USA.

Corresponding author: Warren J. Cheung, MD, MMEd, The Ottawa Hospital, Ottawa, Ontario, Canada, wcheung@toh.ca

Received September 9, 2024; revisions received March 13, 2025, and May 21, 2025; accepted May 22, 2025.