A Critical Disconnect: Residency Selection Factors Lack Correlation With Intern Performance

John C. Burkhardt, MD, PhD
Kendra P. Parekh, MD
Fiona E. Gallahue, MD
Kory S. London, MD
Mary A. Edens, MD
A. J. Humbert, MD
M. Tyson Pillow, MD
Sally A. Santen, MD, PhD
Laura R. Hopson, MD

ABSTRACT

Background Emergency medicine (EM) residency programs want to employ a selection process that will rank best possible applicants for admission into the specialty.

Objective We tested if application data are associated with resident performance using EM milestone assessments. We hypothesized that a weak correlation would exist between some selection factors and milestone outcomes.

Methods Utilizing data from 5 collaborating residency programs, a secondary analysis was performed on residents trained from 2013 to 2018. Factors in the model were gender, underrepresented in medicine status, United States Medical Licensing Examination Step 1 and 2 Clinical Knowledge (CK), Alpha Omega Alpha (AOA), grades (EM, medicine, surgery, pediatrics), advanced degree, Standardized Letter of Evaluation global assessment, rank list position, and controls for year assessed and program. The primary outcomes were milestone level achieved in the core competencies. Multivariate linear regression models were fitted for each of the 23 competencies with comparisons made between each model's results.

Results For the most part, academic performance in medical school (Step 1, 2 CK, grades, AOA) was not associated with residency clinical performance on milestones. Isolated correlations were found between specific milestones (eg, higher surgical grade increased wound care score), but most had no correlation with residency performance.

Conclusions Our study did not find consistent, meaningful correlations between the most common selection factors and milestones at any point in training. This may indicate our current selection process cannot consistently identify the medical students who are most likely to be high performers as residents.

Introduction

The current residency selection process is a time-consuming, expensive venture for training programs and their departments.[1] While training programs actively seek applicants who will succeed and thrive in residency, they also attempt to identify and avoid applicants who will require significant, dedicated, time-consuming resources to fulfill the minimum clinical and professional competency standards. Determining factors that are associated with thriving (or struggling) through training is so far an enigma, but still merit further investigation.

Residency program directors have long considered which metrics to use in an attempt to make reasoned selection decisions.[2,3] These metrics often include standardized testing, clinical grades, and a residency interview process.[3–8] These may become more important as there are changes in Step 1 scoring to pass/fail and the "pause" in the United States Medical Licensing Examination (USMLE) for clinical skills. Unfortunately, the predictive value for success in residency is generally very low based on these metrics, with the majority of positive correlations between standardized testing outcomes.[7,9–14]

The limitations of the interview process have led to the development of alternative interview approaches and the search for other "noncognitive" applicant assessments.[15–19]

One of these assessments is the Standardized Letter of Evaluation (SLOE) used in emergency medicine (EM).[20] The SLOE provides a specialty-based norm-referenced global assessment of each student and their projected location on the program's match rank list.[21] Other methods used to measure noncognitive aspects of applicant bias include the use of novel interview techniques,[18,22] standardized letters of evaluation,[23] and placing less weight on standardized examinations to reduce racial bias.[24] Yet it is clear that bias continues to exist in each step of the selection process.[25] All of these factors increase the priority of ensuring a selection process that is rapid, equitable, and reliable at selecting candidates who will be successful in residency.

Though the challenge of resident selection and metrics for selection has been a topic of repeated

DOI: http://dx.doi.org/10.4300/JGME-D-20-00013.1

Editor's Note: The online version of this article contains regression models with coefficients and competency scores.

research, residency selection factors have a largely unknown predictive correlation with training outcomes or have demonstrated poor predictive value.[11,26–28] The research has been limited by single institution or single training program studies,[29] conducted in a single homogeneous region,[30] limited to just a few core competencies,[19] conducted over a short time period with few residents, or lacked standard outcome measures (ie, varying definitions and measures of success).[31]

The development of milestone assessments has provided a potential solution to the issue of nonstandardized residency outcomes. Milestones as assessment tools were developed "from a close collaboration among the ABMS certifying boards, the review committees, medical-specialty organizations, program-director associations, and residents… to provide meaningful data on the performance that graduates must achieve before entering unsupervised practice."[32] In addition to the milestones' proposed benefits in enhancing residency education quality, patient safety, and driving innovation in graduate medical education, they were also designed to allow for "comparative data" across residency programs.[32] The milestone assessment process has continued to undergo revision, reiteration, and validation to better represent the specific needs of each medical specialty.[33–36]

The objective of this study is to explore whether application and selection factors predict residents' performance in residency at the conclusion of the postgraduate year (PGY) 1 year. We reviewed factors utilized commonly in selection decisions as well as those factors previously identified to be predictive of success or remediation.[19,29,30,37–39] As the milestone assessment was designed to provide for a standard generalizable outcome for residency performance across graduate medical education programs in the same specialty we have used them as our outcomes in this study.

What was known and gap
Residency programs want to employ a selection process that will rank best possible applicants for admission into the specialty. Residency selection factors have a largely unknown predictive correlation with training outcomes or have demonstrated poor predictive value.

What is new
A secondary analysis on residents trained from 2013 to 2018 in 5 residency programs to look for correlation between some program selection factors and milestone outcomes.

Limitations
Stability of program leadership was not assessed, and it is possible intra-program differences in assessments were due to differences in assessors rather than resident specific attributes.

Bottom line
The study did not find consistent, meaningful correlations between the most common selection factors and residency milestones at any point in training.

Methods

Setting and Participants

The study uses secondary data from 5 EM residency programs. The combined dataset included all residents from the entering intern classes from 2010 to 2018. As the EM Milestones were first published in 2012, the outcome data ranges from academic years 2013 to 2018. The EM Milestones were updated in 2015, but there were no substantive changes to the prompts and no changes to the actual milestones aside from their order listed.

Outcomes Measured

Selection factors in the model were gender, underrepresented in medicine (UiM) status; USMLE Step 1 and 2 CK; AOA awards, grades in EM, medicine, surgery, and pediatrics; advanced degree; SLOE global assessment; and rank list position. These were selected based on the available literature about what EM residency program directors used in their decision-making process as well as demographic identities that may correlate with bias in applicant ranking. Gender, UiM status, and AOA awarded were measured as binary factors. Standardized coefficients were used in the calculations and reporting of results for all continuous measures. Models were initially created that included clerkship grades received as categorical variables. These were compared with alternative models that treated these same variables as continuous with each categorical shift treated as an increase of 1 point. Ultimately, the latter was utilized to make comparisons across so many models and variables feasible. As was done with similar studies in EM in the past,[40] interview scores were considered but were not thought to be generalizable by program since each program uses different processes and scoring rubrics for interviews. Rank list is believed to correlate with interview performance; however, given the inclusion of the other factors also thought to correlate with rank list (grades, step scores, etc), it represents unaccounted for decision-making made by program directors based on interview performance and other non-recorded factors.[41] Controls for training start date and specific residency program were also included. A variable for PGY-1–PGY-3 vs PGY-1–PGY-4 format was considered but was dropped as it was colinear with the individual residency program identifier. The primary outcomes of study were milestone level achieved in each of the core competencies after year 1 of training. Milestones were measured in 0.5 increments, which allowed for

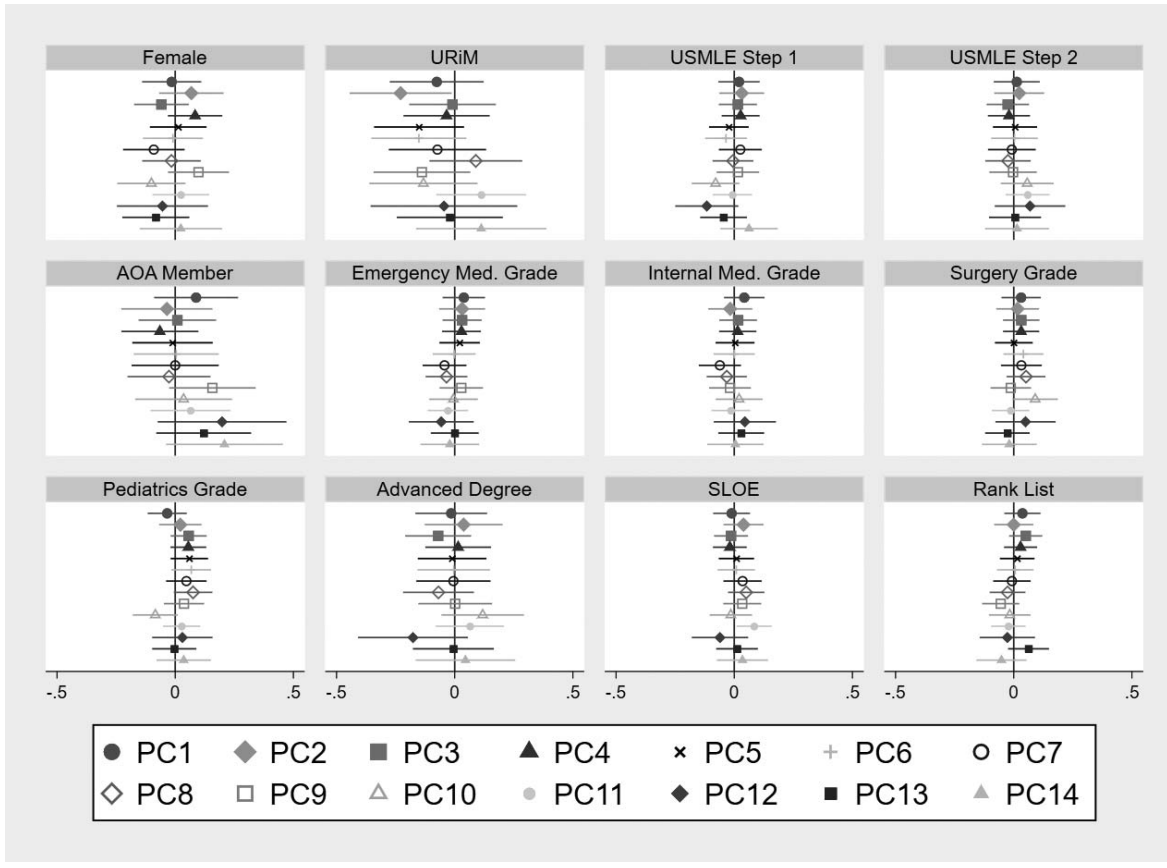

FIGURE 1
Year One Patient Care Competencies

scoring in-between the competency anchor statements. Given the fluidity of scoring between anchoring categories the outcome core competencies were treated as a continuous variable for analysis.

Analysis of the Outcomes

Multivariate linear regression models were fitted for each of the 23 competencies (TABLE 1) with comparisons made between each model's results. This resulted in a total of 23 regression models used in this study. Each individual variable's coefficients for all core competencies were divided into 6 regression coefficient plots (patient care and non–patient care core competency by year).[42] Given the multiple comparisons,[43] a Bonferroni correction for family-wise error rate and a Benjamini–Hochberg false discovery rate were also calculated and provided for comparison as a more conservative estimate of potential correlation.

Institutional Review Board review was solicited at the primary site (where the centralized databased was housed and statistical analysis performed) and all other participating residency programs. The study was determined to be exempt from further review in all cases. Data use agreements were created between the primary site and all other residency sites for deidentified data transfer.

Results

A full account of the 5 participating residency programs revealed 418 individuals for which demographic data were available. Individuals whose milestone records were not available limited the sample size to 329. Resident subjects with data in all 12 selection variables (plus 2 control variables: training start date and individual residency program identifier) dropped the sample size to 213 (TABLE 2). Demographic information on residents included in the initial study group are found in TABLE 2. Variables including Step 1 and 2 CK, clinical grades, and rank list were not associated with EM residents' performance after the first year of residency (FIGURE 1). Having an advanced degree prior to the onset of residency training had a small negative partial correlation (-0.19, 95% CI -0.34 to -0.05) with ICS1 (Patient Centered Communication, FIGURE 2). SLOE global assessment had a small positive partial correlation (0.08, 95% CI 0.01–0.16) with PC11

TABLE 1
Emergency Medicine Core Competencies

Competency Code	Description of Core Competency
PC1	Emergency Stabilization: Prioritizes critical initial stabilization action and mobilizes hospital support services in the resuscitation of a critically ill or injured patient and reassesses after stabilizing intervention.
PC2	Performance of Focused History and Physical Exam: Abstracts current findings in a patient with multiple chronic medical problems and, when appropriate, compares with a prior medical record and identifies significant differences between the current presentation and past presentations.
PC3	Diagnostic Studies: Applies the results of diagnostic testing based on the probability of disease and the likelihood of test results altering management.
PC4	Diagnosis: Based on all of the available data, narrows and prioritizes the list of weighted differential diagnoses to determine appropriate management.
PC5	Pharmacotherapy: Selects and prescribes appropriate pharmaceutical agents based upon relevant considerations such as mechanism of action, intended effect, financial considerations, possible adverse effects, patient preferences, allergies, potential drug-food and drug-drug interactions, institutional policies, and clinical guidelines; and effectively combines agents and monitors and intervenes in the advent of adverse effects in the ED.
PC6	Observation and Reassessment: Re-evaluates patients undergoing ED observation (and monitoring) and, using appropriate data and resources, determines the differential diagnosis, treatment plan, and disposition.
PC7	Disposition: Establishes and implements a comprehensive disposition plan that uses appropriate consultation resources; patient education regarding diagnosis; treatment plan; medications; and time and location specific disposition instructions.
PC8	Multitasking (Task-switching): Employs task switching in an efficient and timely manner in order to manage the ED.
PC9	General Approach to Procedures: Performs the indicated procedure on all appropriate patients (including those who are uncooperative, at the extremes of age, hemodynamically unstable and those who have multiple co-morbidities, poorly defined anatomy, high risk for pain or procedural complications, sedation requirement), takes steps to avoid potential complications, and recognizes the outcome and/or complications resulting from the procedure.
PC10	Airway Management: Performs airway management on all appropriate patients (including those who are uncooperative, at the extremes of age, hemodynamically unstable and those who have multiple co-morbidities, poorly defined anatomy, high risk for pain or procedural complications, sedation requirement), takes steps to avoid potential complications, and recognizes the outcome and/or complications resulting from the procedure.
PC11	Anesthesia and Acute Pain Management: Provides safe acute pain management, anesthesia, and procedural sedation to patients of all ages regardless of the clinical situation.
PC12	Other Diagnostic and Therapeutic Procedures: Goal-directed Focused Ultrasound (Diagnostic/Procedural): Uses goal-directed focused Ultrasound for the bedside diagnostic evaluation of emergency medical conditions and diagnoses, resuscitation of the acutely ill or injured patient, and procedural guidance.
PC13	Other Diagnostic and Therapeutic Procedures: Wound Management: Assesses and appropriately manages wounds in patients of all ages regardless of the clinical situation.
PC14	Other Diagnostic and Therapeutic Procedures: Vascular Access: Successfully obtains vascular access in patients of all ages regardless of the clinical situation.
MK	Demonstrates appropriate medical knowledge in the care of emergency medicine patients.
SBP1	Patient Safety: Participates in performance improvement to optimize patient safety.
SBP2	Systems-based Management: Participates in strategies to improve healthcare delivery and flow. Demonstrates an awareness of and responsiveness to the larger context and system of health care.
SBP3	Technology: Uses technology to accomplish and document safe healthcare delivery.
PBLI	Practice-based Performance Improvement: Participates in performance improvement to optimize ED function, self-learning, and patient care.
PROF1	Professional values: Demonstrates compassion, integrity, and respect for others as well as adherence to the ethical principles relevant to the practice of medicine.

TABLE 1
Continued.

Competency Code	Description of Core Competency
PROF2	Accountability: Demonstrates accountability to patients, society, profession and self.
ICS1	Patient-centered Communication: Demonstrates interpersonal and communication skills that result in the effective exchange of information and collaboration with patients and their families.
ICS2	Team Management: Leads patient-centered care teams, ensuring effective communication and mutual respect among members of the team.

(Anesthesia and Acute Management) after year 1 (FIGURE 1). USMLE Step 2 had a small positive partial correlation with MK (0.01, 95% CI 0–0.02). No other significant partial correlations were found between the selection criteria and core competencies after year 1.

The results reported above were based on independent analysis for each milestone. This represented the most generous number of potential partial correlations in our dataset. Given that multiple comparisons were made as part of the statistical analysis of each core competency, there does exist an increased possibility of a false positive inference. We utilized a Bonferroni correction based on the 23 milestones assessed in each PGY outcome to obtain a more conservative *P* value necessary for statistical significance given our approach ($P = .002$ from 05/23). Following this correction, the partial correlation between having an advanced degree prior to the onset of residency training and ICS1, SLOE and PC11, and Step 2 and MK no longer reached the level of statistical significance.

All 23 regression models with coefficients (partial correlations) of all variables included are available as online supplemental material. Significant differences in competency scores were also identified between programs and between the specific intern class year studied; however, these were used as controls and were not the focus of the research study (provided as online supplemental material).

Discussion

Virtually none of the traditional metrics used in residency selection correlated with milestone performance in the first year of residency. The only partial correlation that survived using statistical corrections for multiple comparisons was the one between USMLE Step 2 and MK. Of note the absolute effect was small with an increase of a standardized deviation on USMLE Step 2 score resulting in an increase in the MK milestone rating of 0.08 points when all other factors were held constant. As the milestone ratings are generally applied in 0.5 increments, more than a 5 SD change would be required to make a practical score change. While "negative studies" often receive little consideration, the most important findings of this study are not what partial correlations were found between selection factors and milestone outcomes but instead their significant absence. These findings demonstrate the ongoing challenge with resident selection in that there is no single factor which independently predicts success (or failure) in graduate medical education training.

TABLE 2
Demographic Variables

Factor	Initial Representation in Dataset (n = 418), No. (%)	Representation in the Effective Sample (n = 329), No. (%)
Residents from PGY-1–PGY-3 programs	220 (52.5)	77 (44.3)
Residents from PGY-1–PGY-4 programs	199 (47.5)	97 (55.8)
Female	149 (35.7)	63 (36.2)
Underrepresented in medicine	45 (10.8)	18 (10.3)
Advanced degree	79 (18.9)	40 (23.0)
AOA member	59 (17.3)	35 (20.1)
Program 1 contribution	122 (29.1)	39 (22.4)
Program 2 contribution	115 (27.5)	34 (19.5)
Program 3 contribution	40 (9.6)	8 (4.6)
Program 4 contribution	65 (15.5)	35 (20.1)
Program 5 contribution	77 (18.4)	58 (33.3)

Abbreviations: PGY, postgraduate year; AOA, Alpha Omega Alpha.
Note: Effective sample represents all individuals with data in all variables in the model and were thus part of the analysis.

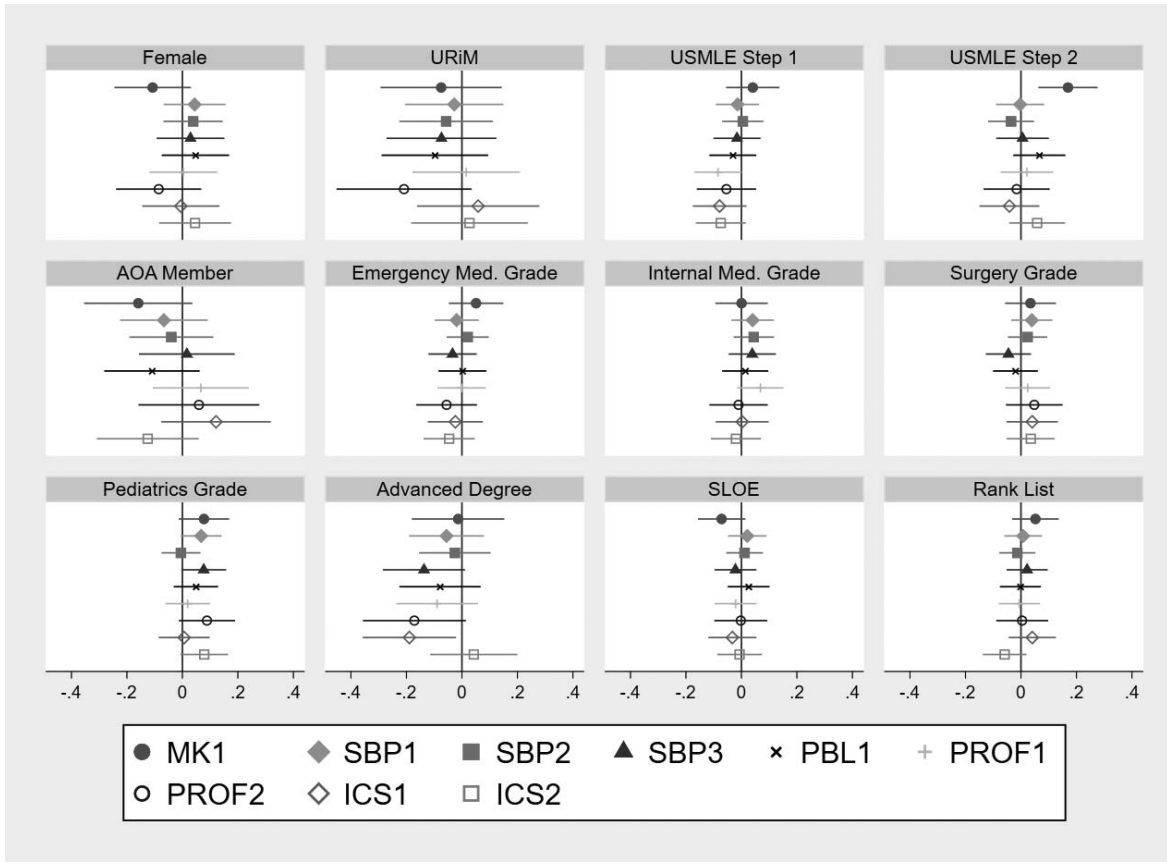

FIGURE 2
Year One Nonpatient Care Competencies

Gender and UiM status remain complex factors in resident selection and future residency success. We included both of these factors in the model to examine our data for signs of bias in scoring. We did not find significant differences between men and women or between UiM and non-UiM trainees in milestone scores within the first year of residency in our data set. In terms of gender differences in other similar studies, a meta-review by Klein et al reported that in 5 of the 9 studies they examined, "a difference in outcomes attributed to gender including gender-based differences in traits ascribed to residents, consistency of feedback, and performance measures" was found.[44] This included articles by Dayal et al, where a significant gender gap in assessments that continued until graduation was seen,[45] Rand et al, where male internal medicine residents scored higher than female residents in 6 of 9 categories,[46] and Mueller et al who found qualitative differences in the content of feedback by attendings to female EM residents.[47] However, a more recent study that incorporated national data from the Accreditation Council for Graduate Medical Education did not find clinically significant differences based on gender.[48]

The interview itself, while not directly included in our study, has also been found to be poorly predictive of training outcomes.[7,12] Residency interviews can also be costly and engender greater bias in selection.[1,12,49–51] As far back as 1979, we can identify an argument by Keck et al that above a certain threshold, traditional cognitive academic criteria have likely reached saturation in predicting those capable of completing medical training and that noncognitive factors such as personality and artistic and social achievement need to be considered.[52] We have not yet identified the "secret sauce" for graduate medical education training success; however, we can continue to strive for residency application metrics that more accurately predict training success and/or more granular measures of residency performance.

Several important limitations exist in this study. First, while this is the largest cohort studying this issue the authors could find and represents programs spread throughout the country, it still represents only a small portion of the population. It is possible that a larger or different cohort could find different partial correlations. While the programs themselves

remained the same, the stability of program leadership was not assessed, and it is possible intra-program differences in assessments were due to differences in assessors rather than resident-specific attributes. Second, while milestone ratings are designed to be a universally applied form of resident outcomes, they are still surrogates for total resident performance and may not be applied in the same way across all residency programs or fully represent the breadth of resident abilities or markers for success. Finally, we conducted multiple comparisons with and without controlling for the potential increase in error generation. When a more conservative standard controlling for multiple comparisons was included (a Bonferroni correction), the significant partial correlations disappeared. In constructing this article, we have included both approaches to provide the most transparent description of how we arrived at our conclusions regarding the lack of predictive accuracy of selection factors on residency outcomes.

Conclusions

Despite efforts to increase standardization of EM clerkship grading and objective assessment of residents with specific measures and prompts, there do not appear to be residency selection factors that partially correlate with resident success during intern year. These findings add to the literature that residency application data which predicts performance in residency remains elusive.

References

1. Van Dermark JT, Wald DA, Corker JR, Reid DG. Financial implications of the emergency medicine interview process. *AEM Educ Train*. 2017;1(1):60–69. doi:10.1002/aet2.10011.
2. Young MJ, Woolliscroft JO, Holloway JJ. Determining the policies of a residency selection committee. *J Med Educ*. 1986;61(10):835–837. doi:10.1097/00001888-198610000-00009.
3. Green M, Jones P, Thomas JX Jr. Selection criteria for residency: results of a national program directors survey. *Acad Med*. 2009;84(3):362–367. doi:10.1097/ACM.0b013e3181970c6b.
4. Berner ES, Brooks CM, Erdmann JB. Use of the USMLE to select residents. *Acad Med*. 1993;68(10):753–759. doi:10.1097/00001888-199310000-00005.
5. Makdisi G, Takeuchi T, Rodriguez J, Rucinski J, Wise L. How we select our residents—a survey of selection criteria in general surgery residents. *J Surg Educ*. 2011;68(1):67–72. doi:10.1016/j.jsurg.2010.10.003.
6. Ruddy MP, Eubanks JEJ, Farrell MEI. More about the role of USMLE Step 1 scores in resident selection. *Acad Med*. 2016;91(11):1468–1469. doi:10.1097/ACM.0000000000001399.
7. Stephenson-Famy A, Houmard BS, Oberoi S, Manyak A, Chiang S, Kim S. Use of the interview in resident candidate selection: a review of the literature. *J Grad Med Educ*. 2015;7(4):539–548. doi:10.4300/JGME-D-14-00236.1.
8. Stringer SP, Cassisi NJ, Slattery WH. Otolaryngology residency selection process: medical student perspective. *Arch Otolaryngol Head Neck Surg*. 1992;118(4):365–366. doi:10.1001/archotol.1992.01880040023004.
9. Liang F, Rudnicki PA, Prince NH, Lipsitz S, May Jr JW, Guo L. An evaluation of plastic surgery resident selection factors. *J Surg Educ*. 2015;72(1):8–15. doi:10.1016/j.jsurg.2014.07.013.
10. McGaghie WC, Cohen ER, Wayne DB. Are United States Medical Licensing Exam Step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? *Acad Med*. 2011;86(1):48–52. doi:10.1097/ACM.0b013e3181ffacdb.
11. Olawaiye A, Yeh J, Withiam-Leitch M. Resident selection process and prediction of clinical performance in an obstetrics and gynecology program. *Teach Learn Med*. 2006;18(4):310–315.
12. Burkhardt JC. What can we learn from resident selection interviews? *J Grad Med Educ*. 2015;7(4):673–675. doi:10.4300/JGME-D-15-00403.1.
13. Swanson DB, Sawhill A, Holtzman KZ, Bucak SD, Morrison C, Hurwitz S, et al. Relationship between performance on part I of the American Board of Orthopaedic Surgery Certifying Examination and scores on USMLE Steps 1 and 2. *Acad Med*. 2009;84(10 Suppl):21–24. doi:10.1097/ACM.0b013e3181b37fd2.
14. Kenny S, McInnes M, Singh V. Associations between residency selection strategies and doctor performance: a meta-analysis. *Med Educ*. 2013;47(8):790–800. doi:10.1111/medu.12234.
15. Eva KW, Reiter HI. Where judgement fails: pitfalls in the selection process for medical personnel. *Adv Health Sci Educ Theory Pract*. 2004;9(2):161–174. doi:10.1023/B:AHSE.0000027479.14637.6c.
16. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ*. 2004;38(3):314–326. doi:10.1046/j.1365-2923.2004.01776.x.
17. Hamel P, Boisjoly H, Corriveau C, Fallaha N, Lahoud S, Luneau K, et al. Using the CanMEDS roles when interviewing for an ophthalmology residency program. *Can J Ophthalmol*. 2007;42(2):299–304.
18. Hopson L, Losman E, Stansfield RB, Vohra T, Turner-Lawerence D, Burkhardt J. The multiple mini interview (MMI) for emergency medicine resident selection.

J Emerg Med. 2014;46:537–543. doi:10.1016/j.jemermed.2013.08.119.
19. Burkhardt JC, Stansfield RB, Vohra T, Losman E, Turner-Lawrence D, Hopson LR. Prognostic value of the multiple mini-interview for emergency medicine residency performance. *J Emerg Med*. 2015;49(2):196–202. doi:10.1016/j.jemermed.2015.02.008.
20. Keim SM, Rein JA, Chisholm C, Dyne PL, Hnedey GW, Jouriles NJ, et al. A standardized letter of recommendation for residency application. *Acad Emerg Med*. 1999;6(11):1141–1146. doi:10.1111/j.1553-2712.1999.tb00117.x.
21. Council of Residency Directors. The Standardized Letter of Evaluation (SLOE). https://www.cordem.org/esloe. Accessed October 7, 2020.
22. Pau A, Jeevaratnam K, Chen YS, Fall AA, Khoo C, Nadarajah VD. The multiple mini-interview (MMI) for student selection in health professions training–a systematic review. *Med Teach*. 2013;35(12):1027–1041. doi:10.3109/0142159X.2013.829912.
23. Li S, Fant AL, McCarthy DM, Miller D, Craig J, Kontrick A. Gender differences in language of standardized letter of evaluation narratives for emergency medicine residency applicants. *AEM Educ Train*. 2017;1(4):334–339. doi:10.1002/aet2.10057.
24. Spector AR, Railey KM. Reducing reliance on test scores reduces racial bias in neurology residency recruitment. *J Natl Med Assoc*. 2019;111(5):471–474. doi:10.1016/j.jnma.2019.03.004.
25. Boatright D, Ross D, O'Connor P, Moore E, Nunez-Smith M. Racial disparities in medical student membership in the Alpha Omega Alpha Honor Society. *JAMA Intern Med*. 2017;177(5):659–665. doi:10.1001/jamainternmed.2016.9623.
26. Hamdy H, Prasad K, Anderson MB, Scherpbier A, Williams R, Zwierstra R, et al. BEME systematic review: predictive values of measurements obtained in medical schools and future performance in medical practice. *Med Teach*. 2006;28(2):103–116. doi:10.1080/01421590600622723.
27. Prober CG, Kolars JC, First LR, Melnick DE. A plea to reassess the role of United States Medical Licensing Examination Step 1 scores in residency selection. *Acad Med*. 2016;91(1):12–15. doi:10.1097/ACM.0000000000000855.
28. Stohl HE, Hueppchen NA, Bienstock JL. Can medical school performance predict residency performance? Resident selection and predictors of successful performance in obstetrics and gynecology. *J Grad Med Educ*. 2010;2(3):322–326. doi:10.4300/JGME-D-09-00101.1.
29. Hayden S, Hayden M, Gamst A. What characteristics of applicants to emergency medicine residency programs predict future success as an emergency medicine resident? *Acad Emerg Med*. 2005;12(3):206–210. doi:10.1197/j.aem.2005.01.002.
30. Van Meter M, Williams M, Banuelos R, Carlson P, Schneider JI, Shy BD, et al. Does the National Resident Match Program rank list predict success in emergency medicine residency programs? *J Emerg Med*. 2017;52(1):77–82.e71. doi:10.1016/j.jemermed.2016.06.059.
31. Wagner JG, Schneberk T, Zobrist M, Hern HG, Jordan J, Boysen-Osboen M, et al. What predicts performance? A multicenter study examining the association between resident performance, rank list position, and United States Medical Licensing Examination Step 1 scores. *J Emerg Med*. 2017;52(3):332–340. doi:10.1016/j.jemermed.2016.11.008.
32. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system—rationale and benefits. *New Engl J Med*. 2012;366(11):1051–1056. doi:10.1056/NEJMsr1200117.
33. Edgar L, Roberts S, Holmboe E. Milestones 2.0: a step forward. *J Grad Med Educ*. 2018;10(3):367–369. doi:10.4300/JGME-D-18-00372.1.
34. Stehman CR, Hochman S, Fernández-Frackelton M, Volz EG, Domingues R, Love JN, et al. Professionalism milestones assessments used by emergency medicine residency programs: a cross-sectional survey. *West J Emerg Med*. 2020;21(1):152–159. doi:10.5811/westjem.2019.11.44456.
35. LaMantia J, Yarris LM, Sunga K, Weizberg M, Hart D, Farina G, et al. Developing and implementing a multisource feedback tool to assess competencies of emergency medicine residents in the United States. *AEM Educ Train*. 2017;1(3):243–249. doi:10.1002/aet2.10043.
36. Beeson MS, Holmboe ES, Korte RC, Nasca TJ, Brigham T, Russ CM, et al. Initial validity analysis of the emergency medicine milestones. *Acad Emerg Med*. 2015;22(7):838–844. doi:10.1111/acem.12697.
37. Bohrer-Clancy J, Lukowski L, Turner L, Staff I, London S. Emergency medicine residency applicant characteristics associated with measured adverse outcomes during residency. *West J Emerg Med*. 2018;19(1):106–111. doi:10.5811/westjem.2017.11.35007.
38. Egol KA, Collins J, Zuckerman JD. Success in orthopaedic training: resident selection and predictors of quality performance. *J Am Acad Orthop Surg*. 2011;19(2):72–80. doi:10.5435/00124635-201102000-00002.
39. Bhat R, Takenaka K, Levine B, Goyal N, Garg M, Visconti A, et al. Predictors of a top performer during emergency medicine residency. *J Emerg Med*. 2015;49(4):505–512. doi:10.1016/j.jemermed.2015.05.035.

40. Breyer MJ, Sadosty A, Biros M. Factors affecting candidate placement on an emergency medicine residency program's rank order list. *West J Emerg Med*. 2012;13(6):458–462. doi:10.5811/westjem.2011.1.6619.
41. Brothers TE, Wetherholt S. Importance of the faculty interview during the resident application process. *J Surg Educ*. 2007;64(6):378–385. doi:10.1016/j.jsurg.2007.05.003.
42. Jann B. Plotting regression coefficients and other estimates. *Stata J*. 2014;14(4):708–737. doi:10.1177/1536867X1401400402.
43. Cao J, Zhang S. Multiple comparison procedures. *JAMA*. 2014;312(5):543–544. doi:10.1001/jama.2014.9440.
44. Klein R, Julian KA, Snyder ED, Koch J, Ufere NN, Volerman A, et al. Gender bias in resident assessment in graduate medical education: review of the literature. *J Gen Intern Med*. 2019;34(5):712–719. doi:10.1007/s11606-019-04884-0.
45. Dayal A, O'Connor DM, Qadri U, Arora VM. Comparison of male vs female resident milestone evaluations by faculty during emergency medicine residency training. *JAMA Intern Med*. 2017;177(5):651–657. doi:10.1001/jamainternmed.2016.9616.
46. Rand VE, Hudes ES, Browner WS, Wachter RM, Avins AL. Effect of evaluator and resident gender on the American Board of Internal Medicine evaluation scores. *J Gen Intern Med*. 1998;13(10):670–674. doi:10.1046/j.1525-1497.1998.00202.x.
47. Mueller AS, Jenkins TM, Osborne M, Dayal A, O'Connor DM, Arora VM. Gender differences in attending physicians' feedback to residents: a qualitative analysis. *J Grad Med Educ*. 2017;9(5):577–585. doi:10.4300/JGME-D-17-00126.1.
48. Santen SA, Yamazaki K, Holmboe ES, Yarris LM, Hamstra SJ. Comparison of male and female resident milestone assessments during emergency medicine residency training: a national study. *Acad Med*. 2020;95(2):263–268. doi:10.1097/ACM.0000000000002988.
49. Eva K, Reiter H, Rosenfield J, Norman G. The ability of the multiple mini-interview to predict preclerkship performance in medical school. *Acad Med*. 2004;79(10):40–42. doi:10.1097/00001888-200410001-00012.
50. Kerfoot BP, Asher KP, McCullough DL. Financial and educational costs of the residency interview process for urology applicants. *Urology*. 2008;71(6):990–994. doi:10.1016/j.urology.2007.11.102.
51. Rosenfeld JM, Reiter HI, Trinh K, Eva KW. A cost efficiency comparison between the multiple mini-interview and traditional admissions interviews. *Adv Health Schi Educ Theory Pract*. 2008;13(1):43–58. doi:10.1007/s10459-006-9029-z.
52. Keck JW, Arnold L, Willoughby L, Calkins V. Efficacy of cognitive/noncognitive measures in predicting resident-physician performance. *J Med Educ*. 1979;54(10):759–765. doi:10.1097/00001888-197910000-00002.

John C. Burkhardt, MD, PhD, is Assistant Professor, Departments of Emergency Medicine and Learning Health Sciences, University of Michigan Medical School; **Kendra P. Parekh, MD,** is Associate Professor, Department of Emergency Medicine, Vanderbilt University School of Medicine; **Fiona E. Gallahue, MD,** is Residency Program Director and Associate Professor, Department of Emergency Medicine, University of Washington; **Kory S. London, MD,** is Associate Residency Program Director, Director of Clinical Operations, Jefferson Methodist ED, Associate Director of Quality Assurance and Practice Improvement, and Assistant Professor, Department of Emergency Medicine, Thomas Jefferson University; **Mary A. Edens, MD,** is Residency Program Director and Associate Professor, Department of Emergency Medicine, Louisiana State University Health Sciences Center Shreveport; **A. J. Humbert, MD,** is Residency Program Director and Associate Professor of Clinical Emergency Medicine, Indiana University School of Medicine; **M. Tyson Pillow, MD,** is Vice Chair of Education, and Associate Professor, Department of Emergency Medicine, Baylor College of Medicine; **Sally A. Santen, MD, PhD,** is Senior Associate Dean for Assessment, Evaluation and Scholarship, and Professor, Department of Emergency Medicine, Virginia Commonwealth University School of Medicine; and **Laura R. Hopson, MD,** is Associate Chair of Education, Emergency Medicine Residency Program, and Associate Professor of Emergency Medicine, University of Michigan Medical School.

Funding: This research was supported in-part by a grant from the University of Michigan Graduate Medical Education Innovation Fund.

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

Corresponding author: John C. Burkhardt, MD, PhD, University of Michigan Medical School, Emergency Medicine and Learning Health Sciences, 1500 East Medical Center Drive, Ann Arbor, MI 48109, 248.563.2557, jburkhar@med.umich.edu

Received January 3, 2020; revisions received May 30, 2020, and July 13, 2020; accepted August 1, 2020.