

Association of Interview and Holistic Review Metrics With Resident Performance-Related Difficulties in an Internal Medicine Residency Program

Reena Hemrajani , MD

Theresa Vettese , MD

Karen Law , MD

Sara Turbow , MD, MPH

ABSTRACT

Background The utility of traditional academic factors to predict residency candidates' performance is unclear. Many programs utilize holistic review processes assessing applicants on an expanded range of application and interview characteristics. Determining which characteristics might predict performance-related difficulty in residency is needed.

Objective We aim to elucidate factors associated with residency performance-related difficulty in a large academic internal medicine residency program.

Methods In 2022, we conducted a retrospective cohort study of Electronic Residency Application Service and interview data for residents matriculating between 2018 and 2020. The primary outcome was a composite of performance-related difficulty during residency (referral to the Clinical Competency Committee; any rotation evaluation score of 2 out of 5 or lower; and/or a confidential "comment of concern" to the program director). Logistic regression models were fit to assess associations between resident characteristics and the composite outcome.

Results Thirty-eight of 117 residents met the composite outcome. Gold Humanism Honor Society (odds ratio [OR] 0.24, 95% confidence interval [CI] 0.16-0.87) or Alpha Omega Alpha (OR 0.36, 95% CI 0.14-0.99) members were less likely to have performance-related difficulty, as were residents with higher United States Medical Licensing Examination Step 2 Clinical Knowledge scores (OR 0.97, 95% CI 0.47-1.00). One-point increases in general faculty overall interview score, leadership competency score, and leadership overall score were associated with 41% to 63% lower odds of meeting the composite outcome. Interview or file review "flags" had an OR of 2.82 (95% CI 1.37-5.80) for the composite outcome.

Conclusions Seven metrics were associated with the composite outcome of resident performance-related difficulty.

Introduction

Selection of candidates into residency programs remains a high-stakes process. As recently as 2021, programs participating in the National Resident Matching Program (NRMP) reported prioritizing traditional selection factors that reflect past academic success, including United States Medical Licensing Examination (USMLE) Step 1 and Step 2 Clinical Knowledge (CK) scores, class ranking, and clerkship scores.¹ Past NRMP and Association of Program Directors in Internal Medicine surveys confirm that internal medicine programs utilize similar academic performance characteristics for interview invitation decisions.^{2,3}

Recently, holistic candidate review, which focuses on applicants' experiences and attributes in addition to academic performance, has been promoted and

increasingly adopted across specialties.^{1,4} While the literature examining resident selection factors and resident performance is extensive, many questions remain. For example, a recent systematic review noted that while Step 2 CK scores predict future in-training and board examination performance, they do not effectively predict subjective performance, with only a weak correlation with overall resident performance, patient care, professionalism, and interpersonal and communication skills.⁵ Similarly, although memberships in Alpha Omega Alpha (AOA) and Gold Humanism Honor Societies (GHHS) were endorsed by 36% and 29%, respectively, of 2021 NRMP program director survey respondents as a consideration in applicant ranking decisions (averaged across all specialties),^{1,6} studies reporting correlation with resident performance or performance-related difficulties in residency are inconclusive.⁷⁻¹⁰ Interviews, while crucial for residency selection,¹¹⁻¹³ have demonstrated mixed results in prediction of performance, problems with professionalism, or attrition.¹⁴⁻²¹

DOI: <http://dx.doi.org/10.4300/JGME-D-22-00726.1>

Editor's Note: The online version of this article contains further data from the study.

In this analysis, we sought to elucidate the relationship between residency selection factors and performance-related difficulties, in the hope that such knowledge may facilitate more thoughtful selection processes and opportunities for enhanced support when at-risk candidates are identified.^{22,23}

Methods

Setting and Participants

This was a retrospective cohort study of all residents matching and matriculating into categorical residency positions at a large academic internal medicine residency program between 2018 and 2020; the study was conducted in 2022.

Applicant Metrics

During the interview process, most applicants interviewed with one “leadership” faculty member (eg, an associate program director, program director, vice chair of education) and one “general” faculty member. Leadership faculty are trained in behavioral-based interviewing (BBI), in which applicants are asked to describe past experiences to facilitate a conversation on traits desirable to the residency program. Each leadership faculty member asks one BBI question during each interview; these questions are the same for an individual faculty member during a single recruitment season and focus on the core areas of teamwork, empathy, and conflict resolution. Leadership faculty give 2 scores: one for the BBI response and one for overall impression. A few residents had 2 leadership interviews due to scheduling variances, rather than one with a faculty member and one with a leadership team member.

General faculty interviewers receive a preparatory article on hiring for emotional intelligence in advance.²⁴ They use an unstructured interview format, asking about topics including, but not limited to, the applicant’s experiences, career interests, and interest in the program. Each general faculty member gives the applicant 4 scores in these competency areas: emotional intelligence, communication skills, career goals, and overall impression.

For both leadership and general faculty interviews, a standardized scoring tool was used to provide a consistent scoring system for each interview conducted. Each competency area was scored from 1 (evidence that a skill or competency is not demonstrated regardless of guidance provided) to 10 (very strong evidence that a skill or competency is present and used effectively). Career goals and overall interview scores also ranged from 1 to 10, with 1 indicating that an applicant should not be ranked, a 6 indicating an applicant who would be a good fit for the program, and 10

KEY POINTS

What Is Known

In the era of increasingly holistic residency application review, additional data is needed that provides correlation between all aspects of the residency application and eventual residency performance.

What Is New

In one internal medicine program, several components of the application review process and interview were associated with fewer performance-related difficulties during residency.

Bottom Line

Residency programs may be interested in devising a similar system to collect data on their applicants’ outcomes in order to better understand who could be at risk for difficulties in their program.

indicating that an applicant should be ranked highly and would be an exceptional resident and excellent fit for the program.

File reviews, including demographics and medical school performance data, were obtained from the Electronic Residency Application Service (ERAS).²⁵ After the interview, associate program directors and core faculty, who previously completed dedicated training on file abstraction and were not involved in the applicant’s interview, abstracted data from each applicant’s ERAS file. Abstracted data included self-identified race/ethnicity, self-identified gender, internal medicine clerkship and internal medicine sub-internship grades, most recent USMLE Step 1 and Step 2 CK scores, and whether the applicant was elected to their school’s chapter of GHHS and/or AOA, if available at their school. GHHS and AOA election were measured as separate variables. Students who did not have GHHS or AOA available at their schools were categorized as GHHS or AOA “not available” for the purposes of this analysis. As clerkship grading scales vary by institution, a simple normalization of scales and aggregation of the normalized data was used to convert clerkship grades across institutions to a 4-point scale modeled on the most common grading scale of Fail/Pass/High Pass/Honors. Race/ethnicity was categorized as a dichotomous variable: underrepresented in medicine (UIM; Black, Latinx, American Indian or Alaskan Native, Native Hawaiian or Pacific Islander) or not UIM. The definition of UIM for this study is based on the AAMC’s 2004 updated definition of the term,²⁶ along with the defined categories provided by ERAS for applicants to self-identify.

Free-text “flags” were noted in the file review or from interviews. Leadership, general faculty interviewers, and file reviewers could identify flags during the interview or file review, respectively. While the criteria for a flag are not explicitly defined, typical

flags reflect concerns such as course or clerkship failure, repeated course or clerkship, USMLE failure, legal issues, portfolio inconsistencies (eg, spelling errors, date inconsistencies, references to residency application for other specialties), or concerns noted in the interview (eg, narrative-application discrepancy, disengaged conversationalist). For this study, each author independently reviewed free-text comments in the flags field for the file review and interview forms to determine whether communicated flags were present and reflected actual concerns. Disagreements or questions were resolved by author consensus. For the analysis, flags were categorized as present/absent, and an individual could have more than one flag identified in their application.

Outcomes Measured

The primary outcome of interest was a composite outcome of performance-related difficulty during residency, defined as meeting any of these criteria: (1) referral to the Clinical Competency Committee (CCC) for a specific competency concern; (2) any score on any rotation evaluation competency-based question of 2 or lower (on a 5-point Likert scale: 1-unacceptable/never, 2-marginal, 3-satisfactory, 4-very good, 5-excellent); or (3) a confidential “comment of concern” submitted to the program director on any evaluation. A composite outcome was examined in order to achieve a large enough sample size for the outcome of performance-related difficulty. The CCC is an Accreditation Council for Graduate Medical Education (ACGME)-mandated committee that reviews resident performance, including specific competency concerns. Rotation evaluation scores were available from 2018 to 2020, for all inpatient and outpatient clinical experiences and included milestone-based questions in all key areas of competency. Peer evaluations were not included as part of the evaluation score component.

Confidential comments, offered as an optional free-text entry on each resident evaluation, were reviewed by the program director and were labeled as positive, neutral, related to wellness, or a comment of concern. Of these, only comments of concern were included as part of the composite outcome. Examples of a comment of concern include additional details regarding deficiency concerns in any of the 6 ACGME core competencies, or behaviors that raised concern for a deficiency in performance that the evaluator might not have known how else to categorize.

Analysis of the Outcomes

Fisher’s exact tests were used to compare categorical variables, and *t* tests and Wilcoxon rank sum tests

were used to compare normally and nonnormally distributed continuous variables, respectively. Unadjusted logistic regression models were created for each file review and interview covariate. Receiver operating characteristic curve values were also reported for each unadjusted model. Covariates that achieved statistical significance were put into adjusted models by category (file review and interview) to determine which file review and interview variables were associated with the outcome of interest; matriculation year was also included in all adjusted models. Correlation between covariates in multivariable models was assessed using the CORRB statement in SAS (SAS Institute Inc). Strongly correlated covariates (correlation >0.4 or <-0.4) were combined or dropped from the model.

This study was approved by the Institutional Review Board of the Emory University School of Medicine and conducted in compliance with ERAS data policy.²⁷

Results

Between 2018 and 2020, 117 residents matched into the residency program. Thirty-eight (32.5%) met the composite outcome of referral to CCC, scoring lower than 2 on any evaluation, and/or having a confidential comment of concern to the program director (TABLE 1).

Of the 38 residents who met the composite outcome, 3 (7.9%) were members of GHHS, compared to 19 of 79 residents (24.0%) who did not meet the composite outcome ($P=.048$; odds ratio [OR] 0.24; 95% confidence interval [CI] 0.16, 0.87; TABLE 1). Residents who met the composite outcome had lower average Step 2 CK scores (247.6 for met outcome vs 253.6 for residents who did not meet the outcome, $P=.04$; OR 0.97; 95% CI 0.94-1.00). Mean faculty overall impression score, leadership behavioral interview score, and leadership overall impression score were all significantly lower for residents who met the criteria for the composite outcome (faculty overall score 7.1 vs 7.8, $P=.02$, OR 0.69, 95% CI 0.47-1.00; leadership behavioral interview score 7.5 vs 8.1, $P=.008$, OR 0.52, 95% CI 0.32-0.87; leadership overall impression score 7.6 vs 8.2, $P=.004$, OR 0.47, 95% CI 0.27-0.81). Residents who met the composite outcome were also more likely to have had one or more flags identified during their interview or file review (18 of 38 [47.3%] of residents who met the composite outcome vs 15 of 79 [19.0%] of residents who did not meet the composite outcome, $P=.003$, OR 2.82, 95% CI 1.37-5.80; TABLE 1). The online supplementary data shows a comparison of descriptive statistics between residents who

TABLE 1

Association Between Resident Characteristics and Performance-Related Difficulty, 2018-2020

Characteristic	Overall	Met Composite Outcome (n=38) ^a	Did Not Meet Composite Outcome (n=79)	P value	OR (95% CI)	ROC
Sex, n (%)						
Male	61 (52.1)	24 (39.3) ^b	37 (60.7)	.07	1.95 (0.88-4.30)	0.58
Female	56 (48.9)	14 (25)	42 (75)			
UIM, n (%)						
Yes	9 (7.7)	5 (55.6)	4 (44.4)	.12	2.84 (0.72-11.26)	0.54
No	108 (92.3)	33 (30.6)	75 (69.4)			
IM Clerkship score, mean (SD) (range 1-4)	3.4 (0.66)	3.3 (0.61)	3.4 (0.68)	.08	0.69 (0.34-1.24)	0.59
IM Sub-internship score, mean (SD) (range 1-4)	3.6 (0.56)	3.4 (0.66)	3.6 (0.49)	.12	0.51 (0.25-1.01)	0.58
GHHS election, n (%)						
Yes	22 (18.8)	3 (13.6)	19 (86.4)	.048[*]	0.24 (0.16-0.87)	0.62
No	75 (61.2)	30 (40)	45 (60)		REF	
Not available	20 (17.1)	5 (25)	15 (75)		–	
AOA member, n (%)						
Yes	36 (30.8)	7 (19.4)	29 (80.6)	.10	0.36 (0.14-0.95)	0.61
No	65 (55.6)	26 (40)	39 (60)		REF	
Not available	16 (13.7)	5 (31.2)	11 (68.7)		–	
USMLE Step 1 score, mean (SD)	240.9 (12.72)	240.6 (15.7)	241.0 (11.8)	.88	1.00 (0.97-1.03)	0.53
Mean USMLE Step 2 CK score, mean (SD)	251.7 (14.17)	247.6 (16.3)	253.6 (12.7)	.04[*]	0.97 (0.94-1.00)	0.61
Overall interview scores						
Faculty interview score, mean (SD) (range 1-10)	7.6 (1.36)	7.1 (1.31)	7.8 (1.34)	.02[*]	0.69 (0.47-1.00)	0.65
Leadership interview score, mean (SD) (range 1-10)	7.9 (0.95)	7.6 (0.86)	8.2 (0.93)	.004[*]	0.47 (0.27-0.81)	0.69
Second leadership interview score, mean (SD) ^c (range 1-10)	7.8 (1.48)	8.0 (–) ^d	7.8 (1.71)	1.00	– ^e	–
Faculty interview scores: individual subcomponents						
Emotional intelligence, mean (SD) (range 1-10)	8.1 (1.16)	8.0 (1.03)	8.2 (1.24)	.34	0.85 (0.56-1.29)	0.56
Communication skills, mean (SD) (range 1-10)	8.1 (1.01)	7.9 (0.92)	8.2 (1.06)	.16	0.77 (0.48-1.25)	0.59
Career goals, mean (SD) (range 1-10)	7.9 (1.03)	8.0 (0.86)	7.9 (1.13)	.39	1.16 (0.73-1.85)	0.56
Leadership interview score: individual subcomponents						
Competency (behavioral) interview mean (SD) (range 1-10)	7.9 (1.04)	7.5 (1.04)	8.1 (0.97)	.008[*]	0.52 (0.32-0.87)	0.67
Competency (behavioral) score: second leadership interview score, mean (SD) (range 1-10)	7.4 (1.82)	7.0 (–) ^d	7.5 (2.08)	1.00	–	–

TABLE 1

Association Between Resident Characteristics and Performance-Related Difficulty, 2018-2020 (continued)

Characteristic	Overall	Met Composite Outcome (n=38) ^a	Did Not Meet Composite Outcome (n=79)	P value	OR (95% CI)	ROC
Flags from either interview or file review (Y/N, n (%))	0: 84 (71.8)	0: 20 (23.8)	64 (76.2)	.003*	2.82 (1.37-5.80)	0.64
	1: 29 (24.8)	1: 16 (55.2)	13 (44.8)			
	2: 3 (2.6)	2: 1 (33.3)	2 (66.7)			
	3: 1 (0.9)	3: 1 (100)	0 (0)			

Abbreviations: OR, odds ratio; ROC, Receiver operating characteristic curve; UIM, underrepresented in medicine; IM, internal medicine; SD, standard deviation; GHHS, Gold Humanism Honor Society; AOA, Alpha Omega Alpha Honor Society; CI, confidence interval; USMLE, United States Medical Licensing Examination; CK, clinical knowledge.

^a The primary composite outcome, performance-related difficulty during residency, was defined as referral to Clinical Competency Committee; any rotation evaluation score ≤ 2 (on a 5-point scale); and/or a confidential "comment of concern" to the program director.

^b Numbers and percentages are calculated across rows, based on Overall.

^c Due to scheduling variations, a small number of candidates participated in 2 separate leadership interviews rather than one leadership and one general faculty interview.

^d Standard deviation could not be calculated due to n=1 in this cell.

^e Odds ratio could not be calculated.

Note: The values that are in bold and have an asterisk are those that are statistically significant results.

met the composite outcome once (n=8) compared to residents who met it more than once (n=30).

TABLE 1 includes ORs for each potential covariate. A 1-point increase in the general faculty interview overall score, leadership competency score, and leadership overall score was associated with between 41% and 63% lower odds of being referred to CCC, receiving a low evaluation, or having a comment of concern to the program director. Identification of any flag on file review or interview was associated with an increase of over 200% in the odds of the composite outcome (OR 2.82, 95% CI 1.37-5.80) for the composite outcome.

TABLE 2 shows the results of the adjusted models, including significant covariates obtained from file review (Step 2 CK score, GHHS election, AOA election); from

interview scores (overall faculty interview score, leadership competency score, leadership overall score); and flags, as well as an indicator variable for matriculation year. The online supplementary data includes the correlation matrix for the covariates in the fully adjusted model; only leadership competency score and leadership overall score were found to have a strong correlation of greater than 0.4 or less than -0.4 (correlation = -0.453), so an average leadership score was created and included in adjusted models. When these 7 covariates were included in a fully adjusted model, both the general faculty overall interview score (adjusted OR [aOR] 0.58, 95% CI 0.35-0.99) and the average leadership score (aOR 0.44, 95% CI 0.20-1.00) remained significant. The results of the adjusted models with the overall leadership score and competency leadership

TABLE 2

Adjusted Logistic Regressions for Composite Outcome^a Among Residents

Application Components	File Review	Interview	Full
GHHS membership (elected vs not elected)	0.29 (0.07-1.12)	–	0.50 (0.08-3.07)
AOA membership (elected vs not elected)	0.51 (0.18-1.47)	–	0.28 (0.05-1.66)
USMLE Step 2 CK score	0.98 (0.95-1.01)	–	0.98 (0.94-1.03)
Leadership score (average of overall + competency score)	–	0.47 (0.24-0.95)	0.44 (0.20-1.00)
Overall score: faculty interview	–	0.69 (0.46-1.04)	0.58 (0.35-0.99)
Flags from interview(s) or file review	–	1.27 (0.51-3.17)	1.38 (0.47-4.08)

Abbreviations: GHHS, Gold Humanism Honor Society; AOA, Alpha Omega Alpha Honor Society; USMLE, United States Medical Licensing Examination; CK, clinical knowledge.

^a The primary composite outcome, performance-related difficulty during residency, was defined as referral to Clinical Competency Committee; any rotation evaluation score ≤ 2 (on a 5-point scale); and/or a confidential "comment of concern" to the program director.

Note: All models include resident matriculation year. GHHS and AOA membership are the odds ratios for elected versus unelected trainees who had GHHS or AOA available at their school.

Note: The values that are in bold are those that are statistically significant results.

score separated are displayed in the online supplementary data; in this model, only the general faculty overall score remained significantly associated with the composite outcome (aOR 0.59, 95% CI 0.35-0.99).

Discussion

We sought to understand associations between residency selection factors and performance-related difficulty during residency. Three residency selection factors from the application: higher Step 2 CK scores, election to AOA, and membership in GHHS were more common in residents who did not experience performance-related difficulty. Higher interview scores, particularly those resulting from behavioral interviews done by experienced faculty, were also associated with lower odds of performance-related difficulty, and flags identified during the interview or on the file review were associated with performance-related difficulty in residency.

Our study adds to previous studies that have demonstrated a correlation between USMLE Step 2 score and resident performance.⁵ Previous literature examining the association between AOA membership and resident performance has been mixed,^{16,28,29} but our results support an association. To our knowledge, a correlation between GHHS membership and less performance-related difficulty has not been previously described. Likewise, our study contributes to the growing literature demonstrating that behavioral-based selection interviews correlate with resident performance; in our case, with performance-related difficulties during training.^{21,30}

Finally, faculty identification of flags during ERAS application review or interview was associated with nearly 3 times the odds of experiencing performance-related difficulty in residency. This suggests that holistic review of applications by faculty may detect subtle areas of concern not otherwise reflected in large-scale review processes, thereby identifying residents who may experience performance-related difficulty, even when scores on standardized, numeric components of the application fall within a program's typical range.

Flag notations by experienced faculty may signal a mismatch between candidates and programs. This may also be reflected in the observed association between general faculty "overall impression" interview scores and resident performance outcomes. Indeed, this study suggests that behavioral-based interviewing and holistic review may be significant tools for identifying residents who may experience performance-related difficulty during training. The results of our study illustrate how investment in faculty development in these key areas supports more effective resident selection, especially when employed in a multifaceted approach of interviews and file reviews.

The goal of identifying residents who are at higher risk for performance-related difficulty during training is not meant to exclude those residents from selection. Rather, in this era of holistic review, this process affords an opportunity to assess the alignment of a single candidate's goals and likelihood of success with the mission, attributes, and resources of a specific program. Furthermore, learners are not static, and difficulties old and new may arise and resolve throughout medical school, residency, and beyond. Many residents who face performance-related difficulty successfully remediate, contribute positively to our programs and communities, and proceed to fulfilling careers.

There are limitations to our study. As this is a single-institution study, our results may not be generalizable. Some applicant characteristics may not be available in our data, such as scores of previous USMLE attempts. Because the study cohort includes residents who matriculated between 2018 and 2020, not all residents had the same amount of time to achieve the outcome; however, we included matriculation year as a covariate in adjusted models. Faculty receive yearly training on interview and evaluation best practices, but the interview and evaluation scores have not been reviewed for interrater reliability. Additionally, we did not attempt to measure resident strengths or factors that might offset identified red flags or mitigate negative evaluations. Further qualitative work should explore how programs that utilize holistic review balance candidate strengths and weaknesses during the application process.

Conclusions

GHHS membership, AOA membership, higher Step 2 CK score, higher behavioral-based interview scores, and absence of red flags through holistic review of residency candidate applications were associated with reduced performance-related difficulty in our program.

References

1. National Resident Matching Program. Results of the 2021 NRMP Program Director Survey. Published August 2021. Accessed August 14, 2023. <https://www.nrmp.org/wp-content/uploads/2021/11/2021-PD-Survey-Report-for-WWW.pdf>
2. Angus SV, Williams CM, Stewart EA, Sweet M, Kisielewski M, Willett LL. Internal medicine residency program directors' screening practices and perceptions about recruitment challenges. *Acad Med*. 2020;95(4):582-589. doi:10.1097/acm.0000000000003086
3. Roberts C, Khanna P, Rigby L, et al. Utility of selection methods for specialist medical training: a BEME (best evidence medical education) systematic review: BEME

- guide no. 45. *Med Teach*. 2018;40(1):3-19. doi:10.1080/0142159X.2017.1367375
4. Association of American Medical Colleges. Holistic review. Accessed June 13, 2022. <https://www.aamc.org/services/member-capacity-building/holistic-review>
 5. Shirkhodaie C, Avila S, Seidel H, Gibbons RD, Arora VM, Farnan JM. The association between USMLE Step 2 Clinical Knowledge scores and residency performance: a systematic review and meta-analysis. *Acad Med*. 2023;98(2):264-273. doi:10.1097/acm.0000000000005061
 6. National Resident Matching Program. Results of the 2020 NRMP Program Director Survey. Published August 2020. Accessed August 14, 2023. <https://www.nrmp.org/wp-content/uploads/2021/08/2020-PD-Survey.pdf>
 7. Burkhardt JC, Parekh KP, Gallahue FE, et al. A critical disconnect: residency selection factors lack correlation with intern performance. *J Grad Med Educ*. 2020;12(6):696-704. doi:10.4300/jgme-d-20-00013.1
 8. Horan DP, Baldwin K, Purtill JJ, Namdari S. Predictors of success in an orthopaedic residency. *JBJS Rev*. 2021;9(6). doi:10.2106/jbjs.Rvw.20.00180
 9. Rosenthal S, Howard B, Schluskel YR, et al. Does medical student membership in the Gold Humanism Honor Society influence selection for residency? *J Surg Educ*. 2009;66(6):308-313. doi:10.1016/j.jsurg.2009.08.002
 10. 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
 11. 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
 12. National Resident Matching Program. Results of the 2016 NRMP Program Director Survey. Published June 2016. Accessed August 14, 2023. <https://www.nrmp.org/wp-content/uploads/2021/07/NRMP-2016-Program-Director-Survey.pdf>
 13. Wagoner NE, Suriano JR, Stoner JA. Factors used by program directors to select residents. *J Med Educ*. 1986;61(1):10-21. doi:10.1097/00001888-198601000-00002
 14. Dubovsky SL, Gendel MH, Dubovsky AN, Levin R, Rosse J, House R. Can admissions interviews predict performance in residency? *Acad Psychiatry*. 2008;32(6):498-503. doi:10.1176/appi.ap.32.6.498
 15. Geissler J, VanHeest A, Tatman P, Gioe T. Aggregate interview method of ranking orthopedic applicants predicts future performance. *Orthopedics*. 2013;36(7):e966-e970. doi:10.3928/01477447-20130624-30
 16. Golden BP, Henschen BL, Liss DT, Kiely SL, Didwania AK. Association between internal medicine residency applicant characteristics and performance on ACGME Milestones during intern year. *J Grad Med Educ*. 2021;13(2):213-222. doi:10.4300/jgme-d-20-00603.1
 17. Lee SH, Phan PH, Desai SV. Evaluation of house staff candidates for program fit: a cohort-based controlled study. *BMC Med Educ*. 2022;22(1):754. doi:10.1186/s12909-022-03801-0
 18. Ozuah PO. Predicting residents' performance: a prospective study. *BMC Med Educ*. 2002;2:7. doi:10.1186/1472-6920-2-7
 19. Shiroma PR, Alarcon RD. Selection factors among international medical graduates and psychiatric residency performance. *Acad Psychiatry*. 2010;34(2):128-131. doi:10.1176/appi.ap.34.2.128
 20. Strand EA, Moore E, Laube DW. Can a structured, behavior-based interview predict future resident success? *Am J Obstet Gynecol*. 2011;204(5):446.e1-446.e13. doi:10.1016/j.ajog.2011.02.019
 21. Stephenson-Famy A, Houmar 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
 22. Brenner AM, Mathai S, Jain S, Mohl PC. Can we predict "problem residents"? *Acad Med*. 2010;85(7):1147-1151. doi:10.1097/ACM.0b013e3181e1a85d
 23. Dupras DM, Edson RS, Halvorsen AJ, Hopkins RH Jr, McDonald FS. "Problem residents": prevalence, problems and remediation in the era of core competencies. *Am J Med*. 2012;125(4):421-425. doi:10.1016/j.amjmed.2011.12.008
 24. Bielaszka-DuVernay C. Hiring for emotional intelligence. *Harvard Business Review*. Published November 19, 2008. Accessed June 13, 2022. <https://hbr.org/2008/11/hiring-for-emotional-intelligence>
 25. Association of American Medical Colleges. Students and residents: applying to residencies with ERAS. Accessed February 16, 2022. <https://students-residents.aamc.org/applying-residencies-eras/applying-residencies-eras>
 26. Association of American Medical Colleges. Underrepresented in medicine definition. Accessed May 11, 2023. <https://www.aamc.org/what-we-do/equity-diversity-inclusion/underrepresented-in-medicine>
 27. Association of American Medical Colleges. AAMC policies regarding the collection, use, and dissemination of resident, intern, fellow, and residency, internship, and fellowship application data. Accessed September 30, 2022. <https://www.aamc.org/media/21456/download>
 28. Gudgel BM, Melson AT, Dvorak J, Ding K, Siatkowski RM. Correlation of ophthalmology residency application characteristics with subsequent performance in residency. *J Acad Ophthalmol* (2017). 2021;13(02):e151-e157. doi:10.1055/s-0041-1733932

29. Raman T, Alrabaa RG, Sood A, Maloof P, Benevenia J, Berberian W. Does residency selection criteria predict performance in orthopaedic surgery residency? *Clin Orthop Relat Res*. 2016;474(4):908-914. doi:10.1007/s11999-015-4317-7
30. Marcus-Blank B, Dahlke JA, Braman JP, et al. Predicting performance of first-year residents: correlations between structured interview, licensure exam, and competency scores in a multi-institutional study. *Acad Med*. 2019;94(3):378-387. doi:10.1097/acm.0000000000002429



All authors are with Emory School of Medicine. **Reena Hemrajani, MD**, is Associate Professor, Department of Medicine, Division of Hospital Medicine; **Theresa Vettese, MD**, is Associate Professor,

Department of Medicine, Division of General Internal Medicine **Karen Law, MD**, is Professor, Department of Medicine, Division of Rheumatology; and **Sara Turbow, MD, MPH**, is Associate Professor, Department of Medicine, Division of General Internal Medicine, and Department of Family & Preventive Medicine, Division of Preventive Medicine.

Funding: The authors report no external funding source for this study.

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

The authors would like to thank Mary Ann Kirkconnell Hall, MPH, for editing assistance.

Corresponding author: Reena Hemrajani, MD, Emory School of Medicine, rhemraj@emory.edu, Twitter @ReenaHem

Received September 30, 2022; revisions received January 16, 2023, and August 2, 2023; accepted August 9, 2023.