

A Systematic Review of Metrics Utilized in the Selection and Prediction of Future Performance of Residents in the United States

Jeremy M. Lipman , MD, MHPE
Colleen Y. Colbert , PhD
Rendell Ashton , MD
Judith French , PhD
Christine Warren, MD, MS

Monica Yepes-Rios , MD
Rachel S. King, JD
S. Beth Bierer , PhD, MEd
Theresa Kline , MLIS, AHIP
James K. Stoller , MS, MD

ABSTRACT

Background Aligning resident and training program attributes is critical. Many programs screen and select residents using assessment tools not grounded in available evidence. This can introduce bias and inappropriate trainee recruitment. Prior reviews of this literature did not include the important lens of diversity, equity, and inclusion (DEI).

Objective This study's objective is to summarize the evidence linking elements in the Electronic Residency Application Service (ERAS) application with selection and training outcomes, including DEI factors.

Methods A systematic review was conducted on March 30, 2022, concordant with PRISMA guidelines, to identify the data supporting the use of elements contained in ERAS and interviews for residency training programs in the United States. Studies were coded into the topics of research, awards, United States Medical Licensing Examination (USMLE) scores, personal statement, letters of recommendation, medical school transcripts, work and volunteer experiences, medical school demographics, DEI, and presence of additional degrees, as well as the interview.

Results The 2599 identified unique studies were reviewed by 2 authors with conflicts adjudicated by a third. Ultimately, 231 meeting inclusion criteria were included (kappa=0.53).

Conclusions Based on the studies reviewed, low-quality research supports use of the interview, Medical Student Performance Evaluation, personal statement, research productivity, prior experience, and letters of recommendation in resident selection, while USMLE scores, grades, national ranking, attainment of additional degrees, and receipt of awards should have a limited role in this process.

Introduction

Misalignment of graduate medical education (GME) resident and program attributes is associated with poor resident performance, dissatisfaction, and attrition.¹⁻³ However, the resident recruitment process is complicated and opaque.^{4,5} Though best practices for identifying applicants who will meet program expectations during GME training has received attention, selecting optimal candidates and predicting resident performance remains challenging, prompting bilateral dissatisfaction, turnover, and occasional dismissal.^{6,7} Many programs select residents using assessments not grounded in available evidence.⁸ This creates potential for bias and misalignment of candidates with programs, and portends poor defense of these selection strategies if challenged.⁹⁻¹¹

The objective of this study was to critically examine evidence associated with elements of the US

residency application process regarding selection and future performance of matriculants. The intention is that education leaders will use this information to review and update their recruitment practices consistent with the most recent evidence.^{12,13} Systematic review methodology was selected over other approaches to integrative scholarship to comprehensively address our research question, given the goal to “identify, critically appraise, and distill” the existing literature on this topic.^{14,15}

Methods

A search strategy was developed in conjunction with a medical librarian (T.K.) to capture elements of resident selection criteria and educational outcomes. Comprehensive searches were conducted in Ovid MEDLINE, Ovid Embase, ERIC, Web of Science, and the Cochrane Central Register of Controlled Trials on March 30, 2022. A combination of controlled vocabulary and keywords was used along with truncation and adjacency operators. No date, language, or publication type restrictions were used.

DOI: <http://dx.doi.org/10.4300/JGME-D-22-00955.1>

Editor's Note: The online version of this article contains the full search strategy used in the study and the PRISMA summary.

The full search strategy is included in the online supplementary data. Although a health care education-focused systematic review would usually include health professions outside medicine, those studies were not included given the focus on outcomes specific to residents in the United States.

A systematic review was then conducted concordant with PRISMA guidelines using Covidence software.¹⁶ All aspects of the review were performed manually with no computerized automation of review employed. Inclusion criteria were created through an iterative research team consensus to examine studies investigating the alignment of outcomes for US residents with information available through the Electronic Residency Application Service (ERAS) and interviews. All team members participated in publication screening to identify those addressing the research question. Two team members reviewed each work for inclusion, with conflicts adjudicated by a third. Following screening, each included study was again reviewed and coded by 2 researchers based on ERAS application metrics (research, awards, United States Medical Licensing Examination [USMLE] scores, personal statement, letters of recommendation [LORs], medical school transcript, work and volunteer experience, medical school demographics, and presence of additional degrees). An additional code was applied to studies investigating the impact of ERAS elements on diversity, equity, and inclusion (DEI). These were identified either as explicitly stating they were examining DEI or by their investigation of recruiting those underrepresented in medicine (UIM). The studies associated with each metric were then reviewed in detail and a narrative synthesis generated. Most studies investigated multiple domains and thus were included in the review and synthesis of all associated metrics. Interrater reliability was calculated with Cohen's kappa using Covidence.

"Holistic review" is defined here as it is by the Association of American Medical Colleges (AAMC) as "mission-aligned admissions or selection processes that take into consideration applicants' experiences, attributes, and academic metrics as well as the value an applicant would contribute to learning, practice, and teaching."¹⁷

Results

The search returned 3360 abstracts for screening from which 761 duplicates were removed. Of the remaining 2599 abstracts, 2215 were excluded as irrelevant to the study question. A total of 383 full-text articles were reviewed by 2 reviewers with a third review required for 62 of these (50 removed). Overall, 152 were excluded due to misalignment

with study outcome, design, or setting. Ultimately, 231 were included in the final review (online supplementary data).¹⁸ Interrater reliability was moderate with an average Cohen's kappa of 0.53.

Included studies were published between 1978 and 2023. General concepts or multiple specialties were examined in 73 studies (32.7%). Among specialty-specific work, most were in surgical specialties followed by internal medicine, emergency medicine, and radiology (TABLE 1).

USMLE Step 1 and 2 Clinical Knowledge Scores as Criteria

Conclusions regarding the association of Step 1 and 2 Clinical Knowledge (CK) scores with performance metrics are widely mixed. TABLE 2 provides a summary of the associations between USMLE and UIM recruitment, specialty board outcome, in-training examination (ITE) scores, and clinical performance.

Medical school deans have identified that the transition of Step 1 to pass/fail may increase reliance on Step 2 CK to filter applications.⁷¹ Only 3 low-quality studies were identified to support a specific Step 2 CK score cutoff for this purpose. While 225 on Step 2 CK is the highest reported associated with improved ITE or board examination performance, this number is of little value given the yearly variability in mean and passing scores.^{25,26,35}

Medical School Grades as Criteria

While some articles in this review noted an association between medical school grades and resident performance in residency,^{48,72,73} others were equivocal.^{12,74,75} One group of retrospective studies found that clerkship grades were not predictive of clinical performance in residency.^{1,27,36,43-45,54,65,76-80} In contrast, other studies found an association.^{8,33,37,51,53,60,61,81,82} One study examining pediatric intern performance found that a model containing number of clerkship honors grades, LOR strength, medical school ranking, and attainment of a master's degree explained 18% of the variance in residents' performance on Accreditation Council for Graduate Medical Education (ACGME) Milestones at the end of internship,⁶¹ with the remaining variance unexplained by academic variables. Likewise, academic performance in medical school was found to be associated with residency ITE^{27,77,78} and board scores,^{78,81} though the correlation was weak.⁷⁸ Other studies found no such relationship.^{26,65} The evidence regarding the association between medical student academic problems and resident performance is also equivocal. While an association was identified between "red flags" in an emergency medicine clerkship (deficiencies in LORs or written comments from clerkship

TABLE 1
Specialties Represented in the Literature

Specialty	N (%)
Multiple specialties	73 (32.7)
General surgery	36 (16.1)
Orthopedics	20 (9)
Internal medicine	15 (6.7)
Emergency medicine	13 (5.8)
Radiology	12 (5.4)
Anesthesiology	9 (4)
OB/GYN	9 (4)
Otolaryngology	8 (3.6)
Ophthalmology	6 (2.7)
Family medicine	5 (2.2)
Pediatrics	5 (2.2)
Urology	3 (1.3)
Neurosurgery	2 (0.9)
Plastic surgery	2 (0.9)
Radiation oncology	2 (0.9)
Dermatology	1 (0.4)
Pathology	1 (0.4)
PM&R	1 (0.4)

Abbreviations: OB/GYN, obstetrics and gynecology; PM&R, physical medicine and rehabilitation.

rotations) and negative outcomes in residency,⁵² other studies found no significant associations between problematic outcomes in residency and medical school academic performance.^{3,38,45,83}

Notably, most studies examining grades as predictor variables were carried out at single institutions.^{2,27,33,36,43,44,50,51,53,61,65,77-82,84,85} As outcome measures differed across studies, results may not be generalizable.⁵¹ In addition, resident performance was often defined subjectively and determined at the end of residency,^{37,60,78} undermining the predictive capability of grades. At least 2 studies cautioned that range restrictions likely affected results, given the competitive nature of their programs.^{2,65} Several studies were conducted before ACGME competencies were introduced,^{50,65,77,79-82,86} and thus cannot be easily compared with more recent studies utilizing Milestone assessments as outcomes.⁸⁴

Clerkship grades are frequently used to differentiate residency applicants. Many authors have noted the variability of grading systems^{37,87} and criteria for honors grades,^{88,89} precluding accurate comparison of applicants across medical schools.^{45,87,88,90} In addition, significant variability exists across clerkships within and between institutions.⁹⁰ Concerns regarding the influence of instructor bias on grades has also been noted.^{87,91} One study found that race and ethnicity were significantly associated with core

clerkship grades.⁹¹ Due to inconsistency in grading and grading systems,⁸⁷ clerkship grades may not be a reliable metric for comparison of students across institutions^{53,87,88,90} or offer an unbiased representation of performance.⁹¹

Medical Student Performance Evaluations as Criteria

Most studies examining the Medical Student Performance Evaluation (MSPE) are descriptive and single-institutional.⁹² These demonstrate that inconsistencies remain in how medical schools apply the AAMC's standardized MSPE template when reporting overall medical student performance,^{83,87,93} normative comparisons such as class rank and grading nomograms,⁹³⁻⁹⁵ or appendices.⁹⁴ Furthermore, discourse analysis of MSPE text suggests the presence of bias associated with MSPE authorship,^{96,97} medical school region,⁹⁷ and applicants' demographic characteristics.^{96,97} Reporting of clerkship grades in MSPEs is more consistent across medical schools in retrospective studies^{93,95} as is accuracy of Alpha Omega Alpha (AOA) awards.⁹⁸ However, one report noted that 30% of top 20 *U.S. News & World Report* medical schools did not report grades in MSPEs as compared to 10% of other schools, which may reflect medical schools' transition to competency-based assessment.⁸⁸ The dearth of MSPE literature provides no^{38,65} to low positive correlational evidence^{3,49,62,83,99-107} between MSPE content and downstream resident performance. Possible MSPE predictors of suboptimal performance during residency include remediation and course failures,^{3,83} medical school leave of absence,⁵¹ negative comments in MSPE,^{3,83} and lower-class rank.³ For instance, in a 20-year retrospective case-control study, 40 psychiatry residents with performance or professionalism concerns during and post residency were included. Of these, 30 were classified as having minor issues where their performance fell below program performance standards but successfully remediated, and 10 residents/graduates classified as having major issues requiring severe program or external governing body action. When compared to 42 matched controls, the 40 who underperformed had more negative MSPE comments, especially the 10 with major performance deficits.⁸³ Total number of clerkship honors reported in MSPE provided low, positive correlational evidence for chief residency status.⁵¹ Another retrospective study of anesthesiology residents showed weak, positive correlations between medical school class rank and satisfactory clinical performance, passing ITEs, publishing one peer-reviewed article, and entering academic practice.³⁷ Importantly, the extent to which medical schools underreport the weaknesses of

TABLE 2

Summary of Data Regarding USMLE 1 and 2 CK as Selection Criteria

Metric	Association Between USMLE Step 1 and 2 CK Scores and Metric	No Association Between USMLE Step 1 and 2 CK Scores and Metric	Mixed Association Between USMLE Step 1 and 2 CK Scores and Metric
UIM applicant recruitment	- An analysis of over 45 000 examinees from 172 medical schools found that female, non-White, and non-native English-speaking test takers score lower on Step 1. Male, non-White, and non-native English speakers score lower on Step 2.¹⁹ - Failure on Step 1 is correlated with being Black, Hispanic, older, female or a first-generation college graduate on review of applications from 6 Midwestern medical schools.²⁰ - UIM and female applicants score lower than others on Step 1. Using Step scores as a screening tool selects against UIM applicants (single and multi-institutional studies).²¹⁻²⁴	No studies identified	No studies identified
Specialty board performance	- Failing Step 1 on the first attempt is correlated with not becoming specialty board certified across many specialties by review of applications from 6 Midwestern medical schools.²⁰ - Smaller studies, mostly single-center, largely found that Step 1 and 2 performance correlates with passing specialty board examinations.²⁵⁻³¹ - A multicentered study identified a positive correlation between failing Step 1 and Orthopedics Board Examination performance.³²	Step 1 scores were not associated with passing the Orthopedics Board Examination in several single-center studies.^{2,33,34}	No studies identified
ITE performance	- Poor performance on Step 1 (any failing or <200 score) is associated with poor performance on ITEs in internal medicine, emergency medicine, orthopedic surgery, anesthesiology, first year of family medicine, urology, and OB/GYN. - Performance on Step 2 is associated with ITE performance in OB/GYN, first year of family medicine, general surgery, emergency medicine, and anesthesiology in single and multi-center studies.^{28,34-41}	One study found no association between Step 1 scores and ITE performance in orthopedics.⁴²	Mixed findings of associations between Step 1 scores and ITE performance in general surgery were identified in numerous single and multi-center studies.^{2,13,27,28,34-38,40,41,43-47}
Clinical performance	A prior systematic review and meta-analysis found that Step 1 and Step 2 CK scores correlated with several performance measures, including summative evaluations, ITEs, and professionalism assessments.⁴⁸	No correlation was found between Step 1 score and clinical performance in internal medicine, orthopedics, OB/GYN, ENT, psychiatry, urology, or pediatrics in single and multi-center studies.^{27,34,36,42,43,47,54-59}	The association of Step 1 score with ACGME Milestones is mixed, with some studies identifying some correlation and others not.^{7,25,51,60-64}

TABLE 2

Summary of Data Regarding USMLE 1 and 2 CK as Selection Criteria (continued)

Metric	Association Between USMLE Step 1 and 2 CK Scores and Metric	No Association Between USMLE Step 1 and 2 CK Scores and Metric	Mixed Association Between USMLE Step 1 and 2 CK Scores and Metric
	▪ Another systematic review found a correlation between Step 1 score and internship supervisor rating across specialties.¹³ ▪ A narrative review of the literature identified that first attempt fail of Step 1 or 2 is correlated with lower clinical performance during residency.⁴⁹ ▪ Across all specialties, 2 years of graduates from a single medical school demonstrated a correlation between higher Step 1 and Step 2 CK scores and top performance ratings during their internship.⁵⁰ ▪ Failing or low Step 1 score was associated with requiring corrective action during emergency medicine and radiology residencies in 2 single-center studies.^{51,52} ▪ Higher Step 2 CK score correlated with better clinical performance in internal medicine from a single residency program and single medical school.^{27,45} ▪ One single-center radiology study found that Step 1 scores were associated with clinical performance.⁵³	▪ A critical review of the literature across specialties found no correlation of Step 1 and Step 2 CK scores with reliable measures of clinical skill acquisition.⁵⁸ ▪ Across multiple specialties, there was no correlation between Step 1 or Step 2 CK scores and selection as a "chief resident" across 13 programs.⁵⁹	▪ Low Step 1 and Step 2 CK scores were associated with poor clinical performance in general surgery in one study, though not in another—both were single-center.^{27,65}
Other criteria	▪ Step 1 scores were found to bias faculty to rate an interview in alignment with the Step 1 score in a single-center study.⁶⁶ ▪ They may also overly influence application reviews from a multi-center study.⁶⁷ ▪ Those elected to the Gold Humanism Honor Society have higher Step 1 scores than those not elected in a 10-institution study.⁶⁸ ▪ Those with higher Step 1 scores were more than 7 times more likely to have inaccurately listed works of scholarship on their ERAS application in a single-center study.⁶⁹ ▪ Step 1 score was inversely correlated with resident awards in surgery at the time of residency graduation in a single-center study.⁷⁰	▪ No studies identified	▪ No studies identified

Abbreviations: USMLE, United States Medical Licensing Examination; CK, clinical knowledge; UIM, underrepresented in medicine; ITE, in-training examination; OB/GYN, obstetrics and gynecology; ENT, ear, nose, and throat; ACGME, Accreditation Council for Graduate Medical Education; ERAS, Electronic Residency Application Service.

their graduates is unknown. An older study identified a 34% prevalence of underreporting events such as leaves of absence and failing grades in MSPEs as compared to school transcripts.⁹⁹

Letters of Recommendation as Criteria

A recent study suggests that structured LORs and standardized letters of evaluation provide more objective and actionable information than traditional narrative LORs.⁴⁹ The structured letters also show improved interrater agreement among readers, and wider use across grading categories, thus enhancing their discriminating power.^{8,100}

LORs are inherently subjective and therefore subject to bias. Many studies have examined whether LORs are systematically biased based on gender,^{101,102} UIM status, or other criteria with mixed results. Some studies show no gender bias, while others show bias toward male applicants and others toward female applicants. There is more consistent evidence for bias against UIM applicants in LORs.¹⁰³

There is little evidence that LORs predict success in training or subsequent practice, except in limited ways. The strongest evidence for the predictive value of the LORs regards the professionalism and humanistic characteristics of applicants.⁵⁴ Compared with standardized test scores and medical school grades, LORs are better predictors of clinical performance during training.²⁷

Personal Statements as Criteria

Personal statements are generally valued by resident selection committees. Most surveyed program leaders note that personal statements are at least moderately important in selecting who gets an interview, assigning rank order, or for assessment during interviews. However, this review found no studies associating personal statements with outcomes during GME training.^{1,74} Their evaluation shows relatively poor interrater reliability, even between evaluators from the same training program.¹⁰⁵

Program leaders who value personal statements tend to use them to assess communication skills and personality.¹⁰⁷ Brevity, precise language, and original thought are considered favorable attributes. Most believe the personal statement is the appropriate place to explain potentially concerning application elements.¹⁰⁴ Problems with personal statements include deceptive or fabricated information, the opportunity for influence from implicit bias, and plagiarism.¹⁰⁸⁻¹¹⁰

Medical School Ranking or Affiliation as Criteria

Adequate data to support the use of *U.S. News & World Report* medical school ranking in a residency

application screening tool were not identified in this review.¹¹¹ There was mixed evidence surrounding whether this ranking is associated with resident clinical performance. One study of radiology residents found that the perceived prestige of the applicant's medical school did not predict resident performance.⁷⁴ The tier of medical school was also not significantly associated with anesthesiology resident performance on any examination, clinical outcome, likelihood of academic publication, or academic career choice.³⁷ In one retrospective study of 46 otolaryngology graduates, a weak correlation was found between rank (in deciles) of the medical school attended and subjective performance evaluation by clinical faculty.¹¹² The authors speculated that residents who attend top-ranked medical schools were a highly select group and thus could predict future success. They also noted their findings may be hampered by affinity bias because they typically enroll students from their affiliated medical school which is ranked in the top decile. There was no statistically significant difference between average ITE scores among students who attended medical school at the same or a different institution as their orthopedic residency (n=60 residents, 2 programs).⁴⁶

Additional Degrees as Criteria

Few studies have examined whether having an additional advanced degree of any type, other than MD/DO, predicts success during residency. Multivariate analysis did not show an association between advanced degree and higher ratings on multisource assessments, higher ITE score, or odds of passing board examinations.⁴⁵ Having an advanced degree was associated with higher patient communication scores.⁴⁵ In one study, anesthesiology residents with additional degrees performed at similar levels as their peers on most outcomes, but tended to be rated lower on clinical performance.³⁷

Research Experience as Criteria

Previous research experience is a readily quantifiable metric in the ERAS application. However, this review did not find associations between resident performance outcomes and research experiences prior to residency across various specialties.^{37,45,46,63,65} Several studies showed weak to moderate correlations between the number of research publications completed prior to application and those completed during residency.¹¹³⁻¹¹⁵ One manuscript found applicants with more first-author publications prior to residency were more likely to pursue fellowship, have a higher h-index (an author-level metric that measures the productivity and citation impact of publications), and publish more during and after residency.¹¹⁵ This review also identified

several studies finding applicants with publications prior to residency were more likely to pursue an academic career.¹¹⁵⁻¹¹⁷

A large body of research across multiple specialties examined problems with erroneous publications listed on applications. A wide range (1%-45%) of publications listed on applications could not be verified as published or contained inaccuracies such as author order.¹¹⁸⁻¹²⁴

Volunteer and Work Experience as Criteria

A 7-year retrospective cohort study of 110 residents showed no association between volunteerism and clinical performance. However, one study found an association between having a career prior to medical school for at least 2 years and competency in interpersonal and communication skills and systems-based practice.¹²⁵ Excellence in athletics, specifically in a team sport, was associated with otolaryngology faculty assessment of clinical excellence,¹¹² clinical performance and completion of general surgery residency,⁴⁴ and selection as chief resident in radiology. A stronger association was noted among college and elite athletes.⁵¹ In one study of an anesthesiology residency program, a negative association was found between leadership experience and ITE and board examination performance. Also, service experience was associated with lower ITE scores.³⁷

There is a paucity of data regarding the strength of association between personal and professional commitment to service and clinical performance in residency. Prior excellence in a team sport may align with success in training.¹¹² No study was identified that evaluated the association of service to underserved communities, membership in medical school affinity groups, health care, or nonprofit work experience with performance in residency.⁷⁴

Medical School Honors and Awards as Criteria

There is mixed evidence regarding the association between AOA membership and residency clinical performance in multiple specialties. AOA membership was associated with higher faculty-defined clinical performance evaluations in anesthesiology and orthopedics programs^{37,78,126} and selection as a chief resident (OR=6.63, $P=.002$).³³ However, AOA award was not predictive of performance in multiple other specialties.^{1,43,49,50,54,65,74,80,112} A retrospective review of internal medicine applications demonstrated strong associations with selection ($P=.0015$), but not with performance in residency as determined by faculty evaluations.⁷⁹

The association between AOA membership and performance on ACGME Milestones across multiple

specialties is equivocal. Although AOA membership was associated with the top third of resident performers, defined by ACGME competencies in 9 emergency medicine programs,⁶⁰ it was not associated with first year performance in emergency or internal medicine, or with professionalism.^{7,53,84,127} AOA status had a negative correlation with patient care Milestones.⁶¹

Evidence from 2 orthopedics studies suggest an association between AOA and passing or higher scores on the ITE, with conflicting evidence on Board examination outcomes.^{26,33,46} Studies from internal medicine and general surgery suggest an association of AOA with Board examination performance.^{26,81} No relationship was found between AOA and faculty assessment of technical skills in general surgery,⁶⁵ or selection for achievement awards.⁷⁰ As noted below, multiple studies have demonstrated a significant bias against UIM applicants for AOA induction (OR 0.16, 95% CI 0.07-0.37).^{21,22,128-130}

This review found a paucity of evidence related to Gold Humanism Honor Society (GHHS) membership and performance in residency. A prior literature review reported a lack of data regarding its impact on ophthalmology residency selection.⁷⁴ A retrospective review of internal medicine residents found a positive association of GHHS with Milestone performance in medical knowledge.⁸⁴

The Interview as a Criterion

This review found mixed evidence regarding resident interviews as predictors of performance.^{48,131} Of the studies reviewed, 24 of the 25 (96%) articles analyzed data collected during a pre-COVID-19, in-person process. One review that examined the virtual interview experience of residency programs both before and during the COVID-19 pandemic found that faculty and applicant feedback was variable.¹³²

One finding shared by several studies is that structured interviews, in which all applicants are asked the same standardized, job-related questions linked to desired program traits, are more likely to predict resident performance than unstructured, conversational interviews.^{74,133,134} Another finding was that multiple factors can potentially bias interview scores, such as interviewer knowledge of board scores and other academic metrics,⁶⁶ as well as applicant physical appearance.^{67,135} Applicants' attractiveness can bias interview evaluations and invitations, especially for women applicants.¹³⁶ Reported associations between interviews and resident performance are provided in TABLE 3.

Diversity, Equity, and Inclusion

USMLE scores, AOA membership, clinical grades, and LOR were identified to be affected by gender,

TABLE 3

Correlation Between Clinical Performance and Factors Related to Applicant Interviews

Metric	Correlated With Clinical Performance	Not Correlated With Clinical Performance
Interview score	- Final rank list submitted to NRMP had a stronger correlation with interview score than with that attained from the ERAS.¹³⁷ - A retrospective multicenter sample of emergency medicine residents found that interview score was significantly associated with resident performance.⁶⁰ - In an 18-year retrospective review of general surgery residents, the interview predicted successful completion of residency, and correlated with clinical and academic performance.⁴⁴ - For a surgery residency program that utilized a standardized personal characteristics form, faculty assessments correlated most favorably with resident clinical performance.²⁷ - A retrospective review of an academic anesthesiology program found that interview performance positively correlated with all ITE scores and clinical performance. Additionally, applicants who received higher scores on interview performance were more likely to pursue academic jobs following residency.³⁷ - Overall, subjective retrospective ratings of otolaryngology graduates by clinical faculty were weakly correlated with interview score.¹¹²	- A retrospective cohort study of internal medicine residents showed that interview score was not associated with Milestone performance.⁸⁴ - A retrospective review of psychiatry residents showed that those with significant problems during training were not predicted by interview ratings.⁸³ - There were no significant differences between groups in admission interview assessments for psychiatry residency between physicians who later developed evidence of impairment and those who did not.¹³⁸ - OB/GYN residents' interview scores did not correlate with resident performance scores.³⁶ - A retrospective cohort study on the degree of concordance between dermatology residency applicant evaluators after a 20-minute conversation found that there was poor agreement among evaluators in assigning application and interview scores. A positive correlation was seen between interview score, academic data, and final rank rather than clinical performance.¹³¹
Structured interviews	- The AAMC SVI Evaluation,¹³⁹ piloted with emergency medicine residency selection during the ERAS 2018-2020 seasons, was stated to be a reliable, valid assessment of behavioral competencies (interpersonal, communication, and professionalism). - SJTs show promise as a method to assess noncognitive attributes such as empathy, integrity, and teamwork but more research is needed.^{140,141} - SJT score predicted overall ACGME Milestone performance in 21 residency programs, including multiple specialties, and also predicted interpersonal and communication skills, and professionalism competencies.¹⁴¹ - Scores on the MMI may correlate with overall resident performance, but the relationship lacks statistical significance when other traditional selection factors are considered.¹⁴² - An MMI score given to emergency medicine interns in the first month of residency correlated with the resident's overall performance in the first year of residency.¹⁴²	The MMI did not correlate significantly with the outcomes when included in multiple regression with other traditional selection factors, such as medical school performance in clerkships, the standard letter of evaluation, medical school ranking, and USMLE scores in an emergency medicine program.¹⁴²
Bias	- A retrospective analysis from a single residency program (n=260 residents) found that a failure to send a post-interview thank you note was a factor associated with a greater likelihood of a negative residency outcome.⁵² - A prospective study found that interview and board scores were significantly correlated when USMLE scores were provided with the OB/GYN application.⁶⁶	No studies identified

TABLE 3

Correlation Between Clinical Performance and Factors Related to Applicant Interviews (continued)

Metric	Correlated With Clinical Performance	Not Correlated With Clinical Performance
	- Applicants to an internal medicine residency were rated on social skills and professional commitment, and then independently rated on physical attractiveness and neatness. Data suggested that neatness and grooming may have had some effect on the interview evaluations and selection of female applicants.¹³⁶ - A deception study examined the impact of facial attractiveness and obesity on radiology resident selection.⁶⁷ Applicant facial attractiveness strongly predicted favorable ratings by faculty, whereas obesity predicted unfavorable ratings, but to a lesser degree. In another study, overall photo scores were associated with invitation to interview.¹³⁵	

Abbreviations: NRMP, National Resident Matching Program; ERAS, Electronic Residency Application Service; ITE, in-training examination; OB/GYN, obstetrics and gynecology; AAMC, Association of American Medical Colleges; SVI, Standardized Video Interview; SJT, Situational Judgement Tests; ACGME, Accreditation Council for Graduate Medical Education; MMI, multiple mini-interview; USMLE, United States Medical Licensing Examination.

racial, and ethnic bias.^{22,91,97,129,130} Reliance on these metrics was found to reduce the number of UIM individuals selected for residency interviews.^{97,128} Three studies found that holistic review of applications is an effective strategy to reduce bias and increase UIM representation.^{22,143,144} Specific strategies that were reported effective included de-emphasizing USMLE Step 1 scores, AOA membership, and grades. Also, some studies reported that bias was reduced by developing selection criteria that include individual applicant experiences and attributes to supplement academic achievement.^{21,22,67,128,135,143,144}

Assessment of applications is subject to the introduction of reviewer biases and substantially impacts the resident selection process.⁹ Therefore, understanding the role of bias is inextricably interwoven with other factors in resident selection. Several studies recommend implicit bias training for those reviewing residency applications, including training to detect bias in letters of recommendation.^{21,22,135,143,144} Such training is associated with recognizing discrimination, personal awareness of bias, and engagement in equity-promoting behaviors.¹⁴⁴ This review did not find any study that analyzed whether training for reviewers is effective in increasing resident diversity. One study found that personal awareness of implicit bias mitigated its effect in the selection process, even without additional training.¹⁴⁵

Discussion

The findings of this review suggest there is minimal evidence aligning residency performance with USMLE score, grades, *U.S. News & World Report* ranking, attainment of additional degrees, technical skills

assessment, and receipt of awards. As such, these elements may be appropriate for a limited role in the assessment of applicants. The MSPE, personal statement, research productivity, prior experience, and LORs may be incorporated in applicant review, with attention to their known limitations. Interviews should be structured, consistent, and include rater training and bias mitigation.

The best-studied parameter in this review is the interview, although limited by the absence of interview format description in most studies and minimal tracking of resident performance over time. While studies were identified to support the association between interview ratings and resident performance, it is evident that potential for bias is high. Studies reviewed did not examine potential biasing factors other than gender, such as race, ethnicity, marital or parental status, and sexual orientation. It is important to acknowledge and mitigate biases for UIM applicants.¹⁴⁶ Supplemental assessments such as situational judgment tests are valuable and cost-effective but require significant effort and expertise to create.^{140,141}

Holistic review of residency applications represents an effective strategy to reduce bias and increase UIM representation.^{22,143,144} Holistic review allows admissions committees to consider the whole applicant, rather than disproportionately focusing on any one factor. The AAMC recommends a 2-step holistic review process in which a program first identifies the experiences, attributes, and academic metrics that align with its goals and values, and then determines how to measure those they have identified.¹⁷

USMLE Step 1 and 2 CK scores are frequently cited as criteria for resident screening. Although Step 1 is now reported only as pass or fail, some applicants still have

numeric scores on their applications. Given the prior reliance on Step 1 scores, it is likely the numeric score on Step 2 CK will replace Step 1 as a screening metric.

The results of this review should be interpreted in the context of its focus on recruitment and selection practices for US GME training programs. Though ample literature addresses resident recruitment and selection in international settings, the distinctive features of training in the United States informs that focus of this review.¹⁴⁷⁻¹⁵⁰ Additionally, an extensive body of research has been developed on recruitment practices for other health and nonhealth professions. However, these articles were not included due to the addition of many potential confounders.¹⁴⁹⁻¹⁵³

Limitations

A significant limitation of this study was the inability to provide summary statistical analysis of the findings. Given the significant heterogeneity of data, including numerous specialties, institutions, and methodologies, such analysis would not be accurate or meaningful. Further, most studies were single-institution and used small samples making extrapolation of results difficult even when pooled. Future research should include larger, multi-institutional studies that can more effectively examine the association between recruitment metrics and residents' performance outcomes across institutions.

Conclusions

This review provides education leaders a summary of the available literature as they consider resident recruitment practices. Though many studies within this systematic review have examined the strength of association between ERAS application criteria and resident performance outcomes, well-designed research is sparse, and results regarding application criteria are mixed.

References

1. 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
2. Dirschl DR, Dahners LE, Adams GL, Crouch JH, Wilson FC. Correlating selection criteria with subsequent performance as residents. *Clin Orthop Relat Res.* 2002;399(399):265-271. doi:10.1097/00003086-200206000-00034
3. Naylor RA, Reisch JS, Valentine RJ. Factors related to attrition in surgery residency based on application data. *Arch Surg.* 2008;143(7):647-651. doi:10.1001/archsurg.143.7.647
4. Berger JS, Cioletti A. Viewpoint from 2 graduate medical education deans application overload in the residency match process. *J Grad Med Educ.* 2016;8(3):317-321. doi:10.4300/JGME-D-16-00239.1
5. Carek PJ, Anderson KD. Residency selection process and the Match: does anyone believe anybody? *JAMA.* 2001;285(21):2784-2785. doi:10.1001/jama.285.21.2784-JMS0606-5-1
6. Khoushhal Z, Hussain MA, Greco E, et al. Prevalence and causes of attrition among surgical residents: a systematic review and meta-analysis. *JAMA Surg.* 2017;152(3):265-272. doi:10.1001/jamasurg.2016.4086
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. Dirschl DR, Champion ER, Gilliam K. Resident selection and predictors of performance: can we be evidence based? *Clin Orthop Relat Res.* 2006;449:44-49. doi:10.1097/01.blo.0000224036.46721.d6
9. Fuchs JW, Youmans QR. Mitigating bias in the era of virtual residency and fellowship interviews. *J Grad Med Educ.* 2020;12(6):674-677. doi:10.4300/JGME-D-20-00443.1
10. Marbin J, Rosenbluth G, Brim R, Cruz E, Martinez A, McNamara M. Improving diversity in pediatric residency selection: using an equity framework to implement holistic review. *J Grad Med Educ.* 2021;13(2):195-200. doi:10.4300/JGME-D-20-01024.1
11. Friedman AM. Using organizational science to improve the resident selection process: an outsider's perspective. *Am J Med Qual.* 2016;31(5):486-488. doi:10.1177/1062860615615669
12. 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
13. Hamdy H, Prasad K, Anderson MB, 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
14. Maggio LA, Samuel A, Stellrecht E. Systematic reviews in medical education. *J Grad Med Educ.* 2022;14(2):171-175. doi:10.4300/JGME-D-22-00113.1
15. McGaghie WC. Varieties of integrative scholarship: why rules of evidence, criteria, and standards matter. *Acad Med.* 2015;90(3):294-302. doi:10.1097/ACM.0000000000000585
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for

- reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71
17. Association of American Medical Colleges. Holistic review. Accessed May 15, 2022. <https://www.aamc.org/services/member-capacity-building/holistic-review>
 18. PRISMA. PRISMA flow diagram. Accessed August 31, 2022. <http://prisma-statement.org/prismastatement/flowdiagram.aspx?AspxAutoDetectCookieSupport=1>
 19. Rubright JD, Jodoin M, Barone MA. Examining demographics, prior academic performance, and United States Medical Licensing Examination scores. *Acad Med*. 2019;94(3):364-370. doi:10.1097/ACM.0000000000002366
 20. McDougale L, Mavis BE, Jeffe DB, et al. Academic and professional career outcomes of medical school graduates who failed USMLE Step 1 on the first attempt. *Adv Health Sci Educ Theory Pract*. 2013;18(2):279-289. doi:10.1007/s10459-012-9371-2
 21. Dorismond C, Farzal Z, Shah R, Ebert C, Buckmire R. Effect of application screening methods on racial and ethnic diversity in otolaryngology. *Otolaryngol Head Neck Surg*. 2022;166(6):1166-1168. doi: 10.1177/01945998221083281
 22. Jarman BT, Kallies KJ, Joshi ART, et al. Underrepresented minorities are underrepresented among general surgery applicants selected to interview. *J Surg Educ*. 2019;76(6):e15-e23. doi:10.1016/j.jsurg.2019.05.018
 23. Poon S, Nellans K, Crabb RAL, et al. Academic metrics do not explain the underrepresentation of women in orthopaedic training programs. *J Bone Joint Surg Am*. 2019;101(8):e32. doi:10.2106/JBJS.17.01372
 24. Grimm LJ, Redmond RA, Campbell JC, Rosette AS. Gender and racial bias in radiology residency letters of recommendation. *J Am Coll Radiol*. 2020;17(1 Pt A):64-71. doi:10.1016/j.jacr.2019.08.008
 25. Harmouche E, Goyal N, Pinawin A, Nagarwala J, Bhat R. USMLE scores predict success in ABEM initial certification: a multicenter study. *West J Emerg Med*. 2017;18(3):544-549. doi:10.5811/westjem.2016.12.32478
 26. Shellito JL, Osland JS, Helmer SD, Chang FC. American Board of Surgery examinations: can we identify surgery residency applicants and residents who will pass the examinations on the first attempt? *Am J Surg*. 2010;199(2):216-222. doi:10.1016/j.amjsurg.2009.03.006
 27. 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
 28. Guffey RC, Rusin K, Chidiac EJ, Marsh HM. The utility of pre-residency standardized tests for anesthesiology resident selection: the place of United States Medical Licensing Examination scores. *Anesth Analg*. 2011;112(1):201-206. doi:10.1213/ANE.0b013e3181fcafcd
 29. De Virgilio C, Yaghoobian A, Kaji A, et al. Predicting performance on the American Board of Surgery qualifying and certifying examinations: a multi-institutional study. *Arch Surg*. 2010;145(9):852-856. doi:10.1001/archsurg.2010.177
 30. McDonald FS, Jurich D, Duhigg LM, et al. Correlations between the USMLE Step Examinations, American College of Physicians In-Training Examination, and ABIM Internal Medicine Certification Examination. *Acad Med*. 2020;95(9):1388-1395. doi:10.1097/acm.0000000000003382
 31. Kay C, Jackson JL, Frank M. The relationship between internal medicine residency graduate performance on the ABIM certifying examination, yearly In-Service Training Examinations, and the USMLE Step 1 Examination. *Acad Med*. 2015;90(1):100-104. doi:10.1097/ACM.0000000000000500
 32. Dougherty PJ, Walter N, Schilling P, Najibi S, Herkowitz H. Do scores of the USMLE Step 1 and OITE correlate with the ABOS Part I certifying examination?: a multicenter study. *Clin Orthop Relat Res*. 2010;468(10):2797-2802. doi:10.1007/s11999-010-1327-3
 33. Turner NS, Shaughnessy WJ, Berg EJ, Larson DR, Hanssen AD. A quantitative composite scoring tool for orthopaedic residency screening and selection. *Clin Orthop Relat Res*. 2006;449:50-55. doi:10.1097/01.BLO.0000224042.84839.44
 34. Crawford CH 3rd, Nyland J, Roberts CS, Johnson JR. Relationship among United States Medical Licensing Step I, orthopedic in-training, subjective clinical performance evaluations, and American Board of Orthopedic Surgery Examination scores: a 12-year review of an orthopedic surgery residency program. *J Surg Educ*. 2010;67(2):71-78. doi:10.1016/j.jsurg.2009.12.006
 35. Thundiyil JG, Modica RF, Silvestri S, Papa L. Do United States Medical Licensing Examination (USMLE) scores predict in-training test performance for emergency medicine residents? *J Emerg Med*. 2010;38(1):65-69. doi:10.1016/j.jemermed.2008.04.010
 36. Bell JG, Kanellitsas I, Shaffer L. Selection of obstetrics and gynecology residents on the basis of medical school performance. *Am J Obstet Gynecol*. 2002;186(5):1091-1094. doi:10.1067/MOB.2002.121622
 37. Chen F, Arora H, Martinelli SM, et al. The predictive value of pre-recruitment achievement on resident performance in anesthesiology. *J Clin Anesth*. 2017;39:139-144. doi:10.1016/J.JCLINANE.2017.03.052
 38. Busha ME, McMillen B, Greene J, Gibson K, Milnes C, Ziemkowski P. One institution's evaluation of family medicine residency applicant data for academic predictors of success. *BMC Med Educ*. 2021;21(1):84. doi:10.1186/S12909-021-02518-W

39. Patzkowski MS, Hauser JM, Liu M, Herrera GF, Highland KB, Capener DC. Medical school clinical knowledge exam scores, not demographic or other factors, associated with residency in-training exam performance. *Mil Med.* 2023;188(1-2):e388-e391. doi:10.1093/milmed/usab332
40. Spurlock DR Jr, Holden C, Hartranft T. Using United States Medical Licensing Examination (USLME) examination results to predict later in-training examination performance among general surgery residents. *J Surg Educ.* 2010;67(6):452-456. doi:10.1016/j.jsurg.2010.06.010
41. Farkas DT, Nagpal K, Curras E, Shah AK, Cosgrove JM. The use of a surgery-specific written examination in the selection process of surgical residents. *J Surg Educ.* 2012;69(6):807-812. doi:10.1016/j.jsurg.2012.05.011
42. Black KP, Abzug JM, Chinchilli VM. Orthopaedic in-training examination scores: a correlation with USMLE results. *J Bone Joint Surg Am.* 2006;88(3):671-676. doi:10.2106/JBJS.C.01184
43. Grewal SG, Yeung LS, Brandes SB. Predictors of success in a urology residency program. *J Surg Educ.* 2013;70(1):138-143. doi:10.1016/J.JSURG.2012.06.015
44. Alterman DM, Jones TM, Heidel RE, Daley BJ, Goldman MH. The predictive value of general surgery application data for future resident performance. *J Surg Educ.* 2011;68(6):513-518. doi:10.1016/J.JSURG.2011.07.007
45. Sharma A, Schauer DP, Kelleher M, Kinnear B, Sall D, Warm E. USMLE Step 2 CK: best predictor of multimodal performance in an internal medicine residency. *J Grad Med Educ.* 2019;11(4):412-419. doi:10.4300/JGME-D-19-00099.1
46. Carmichael KD, Westmoreland JB, Thomas JA, Patterson RM. Relation of residency selection factors to subsequent orthopaedic in-training examination performance. *South Med J.* 2005;98(5):528-532. doi:10.1097/01.SMJ.0000157560.75496.CB
47. Perez JA, Greer S. Correlation of United States Medical Licensing Examination and internal medicine in-training examination performance. *Adv Health Sci Educ Theory Pract.* 2009;14(5):753-758. doi:10.1007/s10459-009-9158-2
48. 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
49. Hartman ND, Lefebvre CW, Manthey DE. A narrative review of the evidence supporting factors used by residency program directors to select applicants for interviews. *J Grad Med Educ.* 2019;11(3):268-273. doi:10.4300/JGME-D-18-00979.3
50. Yindra KJ, Rosenfeld PS, Donnelly MB. Medical school achievements as predictors of residency performance. *J Med Educ.* 1988;63(5):356-363. doi:10.1097/00001888-198805000-00002
51. Maxfield CM, Grimm LJ. The value of numerical USMLE Step 1 scores in radiology resident selection. *Acad Radiol.* 2020;27(10):1475-1480. doi:10.1016/J.ACRA.2019.08.007
52. 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
53. Agarwal V, Bump GM, Heller MT, et al. Do residency selection factors predict radiology resident performance? *Acad Radiol.* 2018;25(3):397-402. doi:10.1016/J.ACRA.2017.09.020
54. Cullen MW, Reed DA, Halvorsen AJ, et al. Selection criteria for internal medicine residency applicants and professionalism ratings during internship. *Mayo Clin Proc.* 2011;86(3):197-202. doi:10.4065/MCP.2010.0655
55. Bowe SN, Schmalbach CE, Laury AM. The state of the otolaryngology match: a review of applicant trends, "impossible" qualifications, and implications. *Otolaryngol Head Neck Surg.* 2017;156(6):985-990. doi:10.1177/0194599817695804
56. Armstrong A, Alvero R, Nielsen P, et al. Do U.S. Medical Licensure Examination Step 1 scores correlate with Council on Resident Education in Obstetrics and Gynecology In-Training Examination Scores and American Board of Obstetrics and Gynecology Written Examination performance? *Mil Med.* 2007;172(6):640-643. doi:10.7205/MILMED.172.6.640
57. Rifkin WD, Rifkin A. Correlation between housestaff performance on the United States Medical Licensing Examination and standardized patient encounters. *Mt Sinai J Med.* 2005;72(1):47-49.
58. 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
59. Cohen ER, Goldstein JL, Schroedl CJ, Parlapiano N, McGaghie WC, Wayne DB. Are USMLE scores valid measures for chief resident selection? *J Grad Med Educ.* 2020;12(4):441-446. doi:10.4300/JGME-D-19-00782.1
60. Bhat R, Takenaka K, Levine B, 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
61. Gross C, O'Halloran C, Winn AS, et al. Application factors associated with clinical performance during

- pediatric internship. *Acad Pediatr*. 2020;20(7):1007-1012. doi:10.1016/J.ACAP.2020.03.010
62. Neely D, Feinglass J, Wallace WH. Developing a predictive model to assess applicants to an internal medicine residency. *J Grad Med Educ*. 2010;2(1):129-132. doi:10.4300/JGME-D-09-00044.1
 63. Hayek SA, Wickizer AP, Lane SM, et al. Application factors may not be predictors of success among general surgery residents as measured by ACGME Milestones. *J Surg Res*. 2020;253:34-40. doi:10.1016/j.jss.2020.03.029
 64. Tolan AM, Kaji AH, Quach C, Hines OJ, de Virgilio C. The electronic residency application service application can predict accreditation council for graduate medical education competency-based surgical resident performance. *J Surg Educ*. 2010;67(6):444-448. doi:10.1016/j.jsurg.2010.05.002
 65. Papp KK, Polk HC, Richardson JD. The relationship between criteria used to select residents and performance during residency. *Am J Surg*. 1997;173(4):326-329. doi:10.1016/S0002-9610(96)00389-3
 66. Smilen SW, Funai EF, Bianco AT. Residency selection: should interviewers be given applicants' board scores? *Am J Obstet Gynecol*. 2001;184(3):508-513. doi:10.1067/mob.2001.109868
 67. Maxfield CM, Thorpe MP, Desser TS, et al. Bias in radiology resident selection: do we discriminate against the obese and unattractive? *Acad Med*. 2019;94(11):1774-1780. doi:10.1097/ACM.0000000000002813
 68. Specter S, Kahn MJ, Lazarus C, et al. Gold Humanism Honor Society election and academic outcomes: a 10-institution study. *Fam Med*. 2015;47(10):770-775.
 69. Yang GY, Schoenwetter MF, Wagner TD, Donohue KA, Kuettel MR. Misrepresentation of publications among radiation oncology residency applicants. *J Am Coll Radiol*. 2006;3(4):259-264. doi:10.1016/j.jacr.2005.12.001
 70. Mainthia R, Ley MJT, Davidson M, Tarpley JL. Achievement in surgical residency: are objective measures of performance associated with awards received in final years of training? *J Surg Educ*. 2014;71(2):176-181. doi:10.1016/j.jsurg.2013.07.012
 71. Manstein SM, Laikhter E, Kazei DD, Comer CD, Shiah E, Lin SJ. The upcoming pass/fail USMLE Step 1 score reporting: an impact assessment from medical school deans. *Plast Surg (Oakv)*. 2023;31(2):169-176. doi:10.1177/22925503211034838
 72. Schaverien MV. Selection for surgical training: an evidence-based review. *J Surg Educ*. 2016;73(4):721-729. doi:10.1016/J.JSURG.2016.02.007
 73. Dooley JH, Bettin KA, Bettin CC. The current state of the residency match. *Orthop Clin North Am*. 2021;52(1):69-76. doi:10.1016/j.ocl.2020.08.006
 74. Lee AG, Golnik KC, Oetting TA, et al. Re-engineering the resident applicant selection process in ophthalmology: a literature review and recommendations for improvement. *Surv Ophthalmol*. 2008;53(2):164-176. doi:10.1016/J.SURVOPHTHAL.2007.12.007
 75. 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
 76. Lee M, Vermillion M. Comparative values of medical school assessments in the prediction of internship performance. *Med Teach*. 2018;40(12):1287-1292. doi:10.1080/0142159X.2018.1430353
 77. Warrick SS, Crumrine RS. Predictors of success in an anesthesiology residency. *J Med Educ*. 1986;61(7):591-595. doi:10.1097/00001888-198607000-00007
 78. 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
 79. George JM, Young D, Metz EN. Evaluating selected internship candidates and their subsequent performances. *Acad Med*. 1989;64(8):480-482. doi:10.1097/00001888-198908000-00013
 80. Borowitz SM, Saulsbury FT, Wilson WG. Information collected during the residency match process does not predict clinical performance. *Arch Pediatr Adolesc Med*. 2000;154(3):256-260. doi:10.1001/ARCHPEDI.154.3.256
 81. Fine PL, Hayward RA. Do the criteria of resident selection committees predict residents' performances? *Acad Med*. 1995;70(9):834-838.
 82. Amos DE, Massagli TL. Medical school achievements as predictors of performance in a physical medicine and rehabilitation residency. *Acad Med*. 1996;71(6):678-680. doi:10.1097/00001888-199606000-00025
 83. 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
 84. 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
 85. Artino AR Jr, Gilliland WR, Waechter DM, Cruess D, Calloway M, Durning SJ. Does self-reported clinical experience predict performance in medical school and internship? *Med Educ*. 2012;46(2):172-178. doi:10.1111/J.1365-2923.2011.04080.X
 86. Dawson-Saunders B, Paiva REA. The validity of clerkship performance evaluations. *Med Educ*.

- 1986;20(3):240-245. doi:10.1111/J.1365-2923.1986.TB01175.X
87. Lin GL, Nwora C, Warton L. Pass/fail score reporting for USMLE Step 1: an opportunity to redefine the transition to residency together. *Acad Med.* 2020;95(9):1308-1311. doi:10.1097/ACM.0000000000003495
 88. Ramakrishnan D, Van Le-Bucklin K, Saba T, Levenson G, Kim JH, Elfenbein DM. What does honors mean? National analysis of medical school clinical clerkship grading. *J Surg Educ.* 2022;79(1):157-164. doi:10.1016/J.JSURG.2021.08.022
 89. Lipman JM, Schenarts KD. Defining honors in the surgery clerkship. *J Am Coll Surg.* 2016;223(4):665-669. doi:10.1016/j.jamcollsurg.2016.07.008
 90. Takayama H, Grinsell R, Brock D, Foy H, Pellegrini C, Horvath K. Is it appropriate to use core clerkship grades in the selection of residents? *Curr Surg.* 2006;63(6):391-396. doi:10.1016/J.CURSUR.2006.06.012
 91. Low D, Pollack SW, Liao ZC, et al. Racial/ethnic disparities in clinical grading in medical school. *Teach Learn Med.* 2019;31(5):487-496. doi:10.1080/10401334.2019.1597724
 92. Association of American Medical Colleges. Recommendations for Revising the Medical Student Performance Evaluation (MSPE). Published 2017. Accessed May 25, 2022. <https://www.aamc.org/media/23311/download>
 93. Hom J, Richman I, Hall P, et al. The state of medical student performance evaluations: improved transparency or continued obfuscation? *Acad Med.* 2016;91(11):1534-1539. doi:10.1097/ACM.0000000000001034
 94. Boysen Osborn M, Mattson J, Yanuck J, et al. Ranking practice variability in the Medical Student Performance Evaluation: so bad, it's "good." *Acad Med.* 2016;91(11):1540-1545. doi:10.1097/ACM.0000000000001180
 95. Brenner JM, Bird JB, Brenner J, Orner D, Friedman K. Current state of the medical student performance evaluation: a tool for reflection for residency programs. *J Grad Med Educ.* 2021;13(4):576-580. doi:10.4300/JGME-D-20-01373.1
 96. Isaac C, Chertoff J, Lee B, Carnes M. Do students' and authors' genders affect evaluations? A linguistic analysis of Medical Student Performance Evaluations. *Acad Med.* 2011;86(1):59-66. doi:10.1097/ACM.0b013e318200561d
 97. Polanco-Santana JC, Storino A, Souza-Mota L, Gangadharan SP, Kent TS. Ethnic/racial bias in medical school performance evaluation of general surgery residency applicants. *J Surg Educ.* 2021;78(5):1524-1534. doi:10.1016/j.jsurg.2021.02.005
 98. Katz ED, Shockley L, Kass L, et al. Identifying inaccuracies on emergency medicine residency applications. *BMC Med Educ.* 2005;5(1):30. doi:10.1186/1472-6920-5-30
 99. Edmond M, Roberson M, Hasan N. The dishonest dean's letter: an analysis of 532 dean's letters from 99 U.S. medical schools. *Acad Med.* 1999;74(9):1033-1035. doi:10.1097/00001888-199909000-00019
 100. Jackson JS, Bond M, Love JN, Hegarty C. Emergency Medicine Standardized Letter of Evaluation (SLOE): findings from the new electronic SLOE format. *J Grad Med Educ.* 2019;11(2):182-186. doi:10.4300/JGME-D-18-00344.1
 101. Khan S, Kirubarajan A, Shamsheri T, Clayton A, Mehta G. Gender bias in reference letters for residency and academic medicine: a systematic review. *Postgrad Med J.* 2023;99(1170):272-278. doi:10.1136/postgradmedj-2021-140045
 102. Lin F, Oh SK, Gordon LK, Pineles SL, Rosenberg JB, Tsui I. Gender-based differences in letters of recommendation written for ophthalmology residency applicants. *BMC Med Educ.* 2019;19(1):476. doi:10.1186/s12909-019-1910-6
 103. Chapman BV, Rooney MK, Ludmir EB, et al. Linguistic biases in letters of recommendation for radiation oncology residency applicants from 2015 to 2019. *J Cancer Educ.* 2022;37(4):965-972. doi:10.1007/s13187-020-01907-x
 104. Hinkle L, Carlos WG, Burkart KM, McCallister J, Bosslet G. What do program directors value in personal statements? A qualitative analysis. *ATS Sch.* 2020;1(1):44-54. doi:10.34197/ats-scholar.2019-0004OC
 105. White BAA, Sadoski M, Thomas S, Shabahang M. Is the evaluation of the personal statement a reliable component of the general surgery residency application? *J Surg Educ.* 2012;69(3):340-343. doi:10.1016/j.jsurg.2011.12.003
 106. Max BA, Gelfand B, Brooks MR, Beckerly R, Segal S. Have personal statements become impersonal? An evaluation of personal statements in anesthesiology residency applications. *J Clin Anesth.* 2010;22(5):346-351. doi:10.1016/j.jclinane.2009.10.007
 107. Mao RMD, Williams TP, Price A, Colvill KM, Cummins CB, Radhakrishnan RS. Predicting general surgery match outcomes using standardized ranking metrics. *J Surg Res.* 2023;283:817-823. doi:10.1016/j.jss.2022.11.038
 108. Segal S, Gelfand BJ, Hurwitz S, et al. Plagiarism in residency application essays. *Ann Intern Med.* 2010;153(2):112-120. doi:10.7326/0003-4819-153-2-201007200-00007
 109. Grover M, Dharamshi F, Goveia C. Deception by applicants to family practice residencies. *Fam Med.* 2001;33(6):441-446.
 110. Smith CJ, Rodenhauer P, Markert RJ. Gender bias of Ohio physicians in the evaluation of the personal statements of residency applicants. *Acad Med.*

- 1991;66(8):479-481. doi:10.1097/00001888-199108000-00014
111. Morse R, Brooks E, Hines K, Wellington S. Methodology: 2023 best medical schools rankings. *U.S. News & World Report*. Published May 10, 2023. Accessed July 1, 2022. <https://www.usnews.com/education/best-graduate-schools/articles/medical-schools-methodology>
112. Chole RA, Ogden MA. Predictors of future success in otolaryngology residency applicants. *Arch Otolaryngol Head Neck Surg*. 2012;138(8):707-712. doi:10.1001/archoto.2012.1374
113. Wright-Chisem J, Cohn MR, Yang J, Osei D, Kogan M. Do medical students who participate in a research gap year produce more research during residency? *J Am Acad Orthop Surg Glob Res Rev*. 2021;5(5):e21.00061. doi:10.5435/JAAOSGlobal-D-21-00061
114. Goss ML, McNutt S, Bible JE. Does publication history predict future publication output in orthopaedics? *Cureus*. 2021;13(5):e15273. doi:10.7759/cureus.15273
115. Namiri NK, Lee AW, Rios N, et al. Predictive factor of preridency publication on career academic achievement in urologists. *Urol Pract*. 2021;8(3):380-386. doi:10.1097/upj.0000000000000208
116. McClelland IS. Pre-residency peer-reviewed publications are associated with neurosurgery resident choice of academic compared to private practice careers. *J Clin Neurosci*. 2010;17(3):287-289. doi:10.1016/j.jocn.2009.07.098
117. Grimm LJ, Shapiro LM, Singhapricha T, Mazurowski MA, Desser TS, Maxfield CM. Predictors of an academic career on radiology residency applications. *Acad Radiol*. 2014;21(5):685-690. doi:10.1016/j.acra.2013.10.019
118. Wiggins MN. A meta-analysis of studies of publication misrepresentation by applicants to residency and fellowship programs. *Acad Med*. 2010;85(9):1470-1474. doi:10.1097/ACM.0b013e3181e2cf2b
119. Hebert RS, Smith CG, Wright SM. Minimal prevalence of authorship misrepresentation among internal medicine residency applicants: do previous estimates of “misrepresentation” represent insufficient case finding? *Ann Intern Med*. 2003;138(5):390-392. doi:10.7326/0003-4819-138-5-200303040-00008
120. Wiggins MN. Misrepresentation by ophthalmology residency applicants. *Arch Ophthalmol*. 2010;128(7):906-910. doi:10.1001/archophthalmol.2010.123
121. Chung CK, Hernandez-Boussard T, Lee GK. “Phantom” publications among plastic surgery residency applicants. *Ann Plast Surg*. 2012;68(4):391-395. doi:10.1097/SAP.0b013e31823d2c4e
122. Yeh DD, Reynolds JM, Pust GD, et al. Publication inaccuracies listed in general surgery residency training program applications. *J Am Coll Surg*. 2021;233(4):545-553. doi:10.1016/j.jamcollsurg.2021.07.002
123. Kistka HM, Nayeri A, Wang L, Dow J, Chandrasekhar R, Chambless LB. Publication misrepresentation among neurosurgery residency applicants: an increasing problem. *J Neurosurg*. 2016;124(1):193-198. doi:10.3171/2014.12.JNS141990
124. Ishman SL, Smith DF, Skinner ML, et al. Unverifiable publications in otolaryngology residency applications. *Otolaryngol Head Neck Surg*. 2012;147(2):249-255. doi:10.1177/0194599812440662
125. Busha ME, McMillen B, Greene J, Gibson K, Channell A, Ziemkowski P. Can life experiences predict readiness for residency? A family medicine residency’s analysis. *J Med Educ Curric Dev*. 2021;8:23821205211062699. doi:10.1177/23821205211062699
126. Kurian EB, Desai VS, Turner NS, et al. Is grit the new fit? Assessing non-cognitive variables in orthopedic surgery trainees. *J Surg Educ*. 2019;76(4):924-930. doi:10.1016/j.jsurg.2019.01.010
127. Cullen M, Wittich C, Halvorsen A, et al. Characteristics of internal medicine residency applicants and subsequent assessments of professionalism during internship. *J Gen Intern Med*. 2010;25(suppl 3):237.
128. Maldjian PD, Trivedi UK. Does objective scoring of applications for radiology residency affect diversity? *Acad Radiol*. 2022;29(9):1417-1424. doi:10.1016/j.acra.2021.11.005
129. Wijesekera TP, Kim M, Moore EZ, Sorenson O, Ross DA. All other things being equal: exploring racial and gender disparities in medical school honor society induction. *Acad Med*. 2019;94(4):562-569. doi:10.1097/ACM.00000000000002463
130. 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
131. 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
132. Yee JM, Moran S, Chapman T. From beginning to end: a single radiology residency program’s experience with web-based resident recruitment during COVID-19 and a review of the literature. *Acad Radiol*. 2021;28(8):1159-1168. doi:10.1016/j.acra.2021.04.009
133. Villwock JA, Bowe SN, Dunleavy D, Overton BR, Sharma S, Abaza MM. Adding long-term value to the residency selection and assessment process.

- Laryngoscope*. 2020;130(1):65-68. doi:10.1002/lary.27878
134. Gardner AK, Grantcharov T, Dunkin BJ. The science of selection: using best practices from industry to improve success in surgery training. *J Surg Educ*. 2018;75(2):278-285. doi:10.1016/j.jsurg.2017.07.010
 135. Kassam A, Cortez AR, Winer LK, et al. Swipe right for surgical residency: exploring the unconscious bias in resident selection. *Surgery*. 2020;168(4):724-729. doi:10.1016/j.surg.2020.05.029
 136. Boor M, Wartman SA, Reuben DB. Relationship of physical appearance and professional demeanor to interview evaluations and rankings of medical residency applicants. *J Psychol*. 1983;113(1st Half):61-65. doi:10.1080/00223980.1983.9923557
 137. Kamangar F, Davari P, Azari R, et al. The residency interview is still paramount: results of a retrospective cohort study on concordance of dermatology residency applicant evaluators and influence of the applicant interview. *Dermatol Online J*. 2017;23(5):13030/qt7rf0x11c.
 138. Dubovsky SL, Gendel M, Dubovsky AN, Rosse J, Levin R, House R. Do data obtained from admissions interviews and resident evaluations predict later personal and practice problems? *Acad Psychiatry*. 2005;29(5):443-447. doi:10.1176/appi.ap.29.5.443
 139. Bird SB, Hern HG, Blomkalns A, et al. Innovation in residency selection: the AAMC standardized video interview. *Acad Med*. 2019;94(10):1489-1497. doi:10.1097/ACM.00000000000002705
 140. Patterson F, Zibarras L, Ashworth V. Situational judgement tests in medical education and training: research, theory and practice: AMEE guide no. 100. *Med Teach*. 2016;38(1):3-17. doi:10.3109/0142159X.2015.1072619
 141. Cullen MJ, Zhang C, Marcus-Blank B, et al. Improving our ability to predict resident applicant performance: validity evidence for a situational judgment test. *Teach Learn Med*. 2020;32(5):508-521. doi:10.1080/10401334.2020.1760104
 142. 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
 143. Hemal K, Reghunathan M, Newsom M, Davis G, Gosman A. Diversity and inclusion: a review of effective initiatives in surgery. *J Surg Educ*. 2021;78(5):1500-1515. doi:10.1016/j.jsurg.2021.03.010
 144. Ware AD, Flax LW, White MJ. Strategies to enhance diversity, equity, and inclusion in pathology training programs: a comprehensive review of the literature. *Arch Pathol Lab Med*. 2021;145(9):1071-1080. doi:10.5858/arpa.2020-0595-RA
 145. Maxfield CM, Thorpe MP, Desser TS, et al. Awareness of implicit bias mitigates discrimination in radiology resident selection. *Med Educ*. 2020;54(7):637-642. doi:10.1111/medu.14146
 146. Nwora C, Allred DB, Verduzco-Gutierrez M. Mitigating bias in virtual interviews for applicants who are underrepresented in medicine. *J Natl Med Assoc*. 2021;113(1):74-76. doi:10.1016/j.jnma.2020.07.011
 147. Wijnen-Meijer M, Burdick W, Alofs L, Burgers C, ten Cate O. Stages and transitions in medical education around the world: clarifying structures and terminology. *Med Teach*. 2013;35(4):301-307. doi:10.3109/0142159X.2012.746449
 148. Weggemans MM, Dijk van B, Dooijeweert van B, Veenendaal AG, ten Cate O. The postgraduate medical education pathway: an international comparison. *GMS J Med Educ*. 2017;34(5):Doc63. doi:10.3205/zma001140
 149. Burgess RM, Ponton MK, Weber MD. Student recruitment strategies in professional physical therapist education programs. *J Phys Ther Educ*. 2004;18(2):22-30.
 150. Hormann HJ, Maschke P. On the relation between personality and job performance of airline pilots. *Int J Aviat Psychol*. 1996;6(2):171-178. doi:10.1207/s15327108ijap0602_4
 151. Knettle M, Nowacki AS, Hrehocik M, Stoller JK. Matching allied health student need with supply: description of a new index. *J Allied Health*. 2021;50(1):e23-e29.
 152. Tett RP, Jackson DN, Rothstein M. Personality measures as predictors of job performance: a meta-analytic review. *Pers Psychol*. 1991;44(4):703-742. doi:10.1111/j.1744-6570.1991.tb00696.x
 153. Barrick MR, Mount MK. The big five personality dimensions and job performance: a meta-analysis. *Pers Psychol*. 1991;44(1):1-26. doi:10.1111/j.1744-6570.1991.tb00688.x



Jeremy M. Lipman, MD, MHPE, is Professor of Surgery, Director of Graduate Medical Education, and Designated Institutional Official (DIO), Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA; **Colleen Y. Colbert, PhD**, is Professor of Medicine, and Director, Office of Educator and Scholar Development, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA; **Rendell Ashton, MD**, is Associate Professor of Medicine, Director, Pulmonary, Critical Care Fellowship, and Associate DIO, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA; **Judith French, PhD**, is Associate Professor of Surgery, and Vice Chair for Education, Department of General Surgery, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western

Reserve University, Cleveland, Ohio, USA; **Christine Warren, MD, MS**, is Associate Professor of Dermatology and Associate Dean, Admissions and Student Affairs, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA; **Monica Yepes-Rios, MD**, is Associate Professor of Medicine and Assistant Dean, Diversity Equity and Inclusion for Students, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA; **Rachel S. King, JD**, is Director of Educational Equity and Title IX Coordinator, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA; **S. Beth Bierer, PhD, MEd**, is Professor of Medicine and Director, Assessment and Evaluation, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA; **Theresa Kline, MLIS, AHIP**, is Medical Librarian, Cleveland Clinic Lerner College of Medicine of Case Western Reserve

University, Cleveland, Ohio, USA; and **James K. Stoller, MS, MD**, is Professor of Medicine, Senior Associate Dean, and Chairman, Education Institute, Cleveland Clinic and Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, Ohio, USA.

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

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

Corresponding author: Jeremy M. Lipman, MD, MHPE, Cleveland Clinic, Cleveland, Ohio, USA, lipmanj@ccf.org, X @JeremyLipman

Received December 15, 2022; revisions received April 26, 2023, and August 30, 2023; accepted September 5, 2023.