Characteristics Associated with Successful Residency Match in General Surgery

Reagan A. Collins, BA,*† Kaitlyn Nimmer, BA,‡ Salma A. Sheriff, BS,‡ Tania K. Arora, MD,§
Anai N. Kothari, MD, MS,‡ Carrie Cunningham, MD, MPH,* and Callisia N. Clarke, MD, MS‡

Objective: To evaluate characteristics of matched and unmatched general surgery residency (GSR) applicants.

Background: Given the recent change of the United States Medical Licensing Exam Step 1 grading to pass/fail, understanding the factors that influence GSR match success is integral to identifying potential interventions to improve match rates for diverse medical students.

Methods: Retrospective review of GSR National Residency Matching Program (NRMP) applicant and Accreditation Council for Graduate Medical Education (ACGME) active resident data between 2011 and 2022. Data included application characteristics for United States ("US") and "independent" applicants, factors cited by program directors in the interview and ranking process, paths pursued if applicants went unmatched, and racial/ethnic representation.

Results: A total of 9149 US and 3985 independent applicants applied to GSR between 2011 and 2021. Matched *versus* unmatched applicants had higher step 1 scores (US: 236 *vs* 218, $P = 0.005$; independent: 237 *vs* 228, $P = 0.001$), higher step 2 scores (US: 248 *vs* 232, $P = 0.006$; independent: 245 *vs* 234, $P < 0.001$), more likely to belong to alpha omega alpha (US: 17.1% *vs* 1.6%, $P = 0.002$) or to attend a top 40 National Institutes of Health-funded school (US: 31.0% *vs* 19.4%, $P = 0.002$) compared to unmatched applicants. Program directors heavily factored step 1 and step 2 scores, letters of recommendation, interactions with faculty and trainees, and interpersonal skills when interviewing and ranking applicants. The proportion of active general surgery residents *versus* applicants was lower for Asians (12.3% *vs* 20.9%, $P < 0.001$), Black/African American (5.0% *vs* 8.8%, $P < 0.001$), Hispanic/Latino (5.0% *vs* 9.4%, $P = 0.001$), and underrepresented in medicine students (10.3% *vs* 19.1%, $P < 0.001$).

Conclusions: In the pass/fail step 1 era, factors including step 2 score and other subjective metrics may be more heavily weighted in the GSR match process.

Keywords: general surgery, match, race/ethnicity, residency, step 1, step 2

INTRODUCTION

The Association of American Medical Colleges (AAMC) has identified promoting a racially and ethnically diverse physician workforce as a key approach to alleviating health inequities in the United States (US).[1] This is in part due to the greater likelihood of minoritized physicians treating minority patients, practicing in underserved areas, and the benefits both culturally sensitive and racially concordant care have on patient-physician relationships and health outcomes.[2,3] Nonetheless, minority surgeons are direly underrepresented; 2% of surgeons are Black,

From the *Department of Surgery, Massachusetts General Hospital, Boston, MA; †School of Medicine, Texas Tech University Health Sciences Center, Lubbock, TX; ‡Division of Surgical Oncology, Department of Surgery, Medical College of Wisconsin, Milwaukee, WI; and §Division of Surgical Oncology, Department of Surgery, Augusta University at the Medical College of Georgia, Augusta, GA.

Reprints: Callisia N. Clarke, MD, MS, Division of Surgical Oncology, Department of Surgery, Medical College of Wisconsin, 8701 W. Watertown Plank Road, Milwaukee, WI 53226. Email: cnclarke@mcw.edu.

5% are Hispanic/Latino, and 0.1% are American Indian/Alaska Native compared to 13.6%, 18.9%, and 1.3% of the US population, respectively.[4,5]

As the US becomes increasingly racially and ethnically diverse, promoting a diverse surgical workforce is essential to providing culturally competent care. General surgery residencies (GSR) have historically emphasized academic characteristics such as United States Medical Licensing Exam (USMLE) scores, clerkship grades, and alpha omega alpha (AOA) membership as surrogates for future residency success, although these vary based on race/ethnicity, socioeconomic status, and training program.[6–10] In 2020, 90% of residency programs used USMLE step 1 scores as an initial screening tool for applicants.[11–14] When using the mean step 1 score in 2014 as a theoretical cutoff, 74.8% of Hispanic applicants and 84.3% of Black applicants would have been filtered out of the application pool.[14] Thus, the overt reliance on academic metrics is known to lead to bias and limit diversity in residency programs, underscoring the need for a more comprehensive, inclusive approach to evaluating applicants' competency.

The change in reporting of USMLE step 1 scores hopes to prompt residency programs to adopt a more holistic review of applicants to evaluate students' clinical aptitude and suitability. A 2021 survey of GSR program directors identified step 2 score, personal contacts (mentors or academic acquaintances), medical school prestige, and sub-internship performance as metrics that would be more heavily weighted following the scoring change.[7] Congruent with previous work, USMLE step 2 score, honors in clerkships, AOA membership, and number of publications are significant predictors of a successful GSR match.[6,9] Despite 70.9% of program directors disagreeing with the pass/fail scoring change, more than 40% agreed that a holistic review would decrease socioeconomic disparities and promote diversity and inclusion in their programs.[7]

A more comprehensive understanding of factors influencing successfully matching into a GSR is necessary to identify potential interventions to improve diversity, equity, and inclusivity in the selection process. We aimed to evaluate academic and demographic characteristics associated with matching into GSR and assess what paths unmatched applicants pursue.

METHODS

Data Source

This is a retrospective review of general surgery (GS) applicants and residents from the 2011–2012 to 2021–2022 academic years. Data were obtained from the National Residency Matching Program (NRMP) Applicant and Program Director Surveys, the NRMP Charting Outcomes from the Match reports for GSR, and the Accreditation Council for Graduate Medical Education (ACGME) Data Resource Books over the study period. This study was deemed exempt from institutional review board approval.

The NRMP Applicant Survey and Charting Outcomes in the Match reports were accessed for data on matched and unmatched applicants who selected a categorical GSR as their "preferred specialty" in the match. Data from each reporting year were recorded and aggregated for the analysis. Characteristics obtained from the charting outcomes in the match reports included research experience, publications, work or volunteer experience, AOA membership, top 40 National Institutes of Health-funded schools, PhD or graduate degree, and USMLE step 1 and 2 scores. Number of applications submitted, interview offers, interviews attended, programs ranked, and paths pursued by unmatched applicants were obtained from the NRMP Applicant Survey. The likelihood of an unmatched applicant pursuing various paths was reported on a 5-point Likert scale, where 5 corresponded to "extremely likely" and 1 to "not at all likely". Likelihood scores for each path were recorded and averaged over the study period.

NRMP Program Director Surveys were assessed for top factors cited for interviewing and ranking an applicant. For the 2012–2013 academic year, factors were reported on a 5-point Likert scale and excluded from the analysis. Percentages of program directors citing each factor were recorded and averaged over the study period.

Self-reported race/ethnicity data of GS applicants and residents were obtained from the ACGME data resource book over the study period. The underrepresented in medicine (URM) category is based on AAMC designation and includes students who self-identify as American Indian, Black/African American, Hispanic/Latino, or Native Hawaiian/Pacific Islander. Applicants from Doctor of Medicine and Doctor of Osteopathic Medicine medical schools, including US and international medical graduates, were included in the analysis.

Based on reporting from the NRMP, "US applicants" included seniors from US allopathic medical schools. "Independent" applicants included allopathic medical school graduates, US citizen and non-US citizen students, graduates of international, osteopathic, and Canadian medical schools, and graduates of Fifth Pathway programs. Starting in 2020, application and interview reporting metrics for Independent applicants were sub-grouped into "DO" or "Other" for ease of comparing data before and after COVID-19 prompted the change to the virtual interview format. The NRMP applicant data clearly differentiates US *versus* independent applicants, however, data on race/ethnicity is only available as an aggregate for all applicants and all active residents and thus is reported as such.

Data Analysis

Data analysis was conducted using R version 4.2.2 with a statistically significant P value < 0.05. Continuous and categorical variables were summarized as median (interquartile range)

and n (percentage), respectively. Matched and unmatched US and independent applicant characteristics were compared using the Wilcoxon rank sum test. The AAMC and NRMP do not publicly report the race/ethnicity of matched and unmatched applicants, thus, differences in race/ethnicity proportions were assessed using the Wilcoxon rank sum test using data from GS applicants and residents.[15,16] Trends in race/ethnicity proportions of GS applicants and residents were assessed using the Mann-Kendall trend test.

RESULTS

Academic Characteristics

During the study period, a total of 9149 US and 3985 independent applicants applied to GSR positions with a 90.1% and 44.9% match rate, respectively. For both US and independent applicants, matched applicants received more interview offers (US: 18.0 interview offers *vs* 8.0, $P = 0.005$; independent: 8.0 *vs* 3.0, $P = 0.009$), attended more interviews (US: 14.0 interviews attended *vs* 8.0, $P = 0.017$; independent: 7.0 *vs* 3.0, $P = 0.009$), and ranked more programs (US: 14.0 programs ranked *vs* 8.0, $P = 0.020$; independent: 6.2 *vs* 2.7, $P < 0.001$) as compared to unmatched applicants (Table 1). Matched applicants had significantly higher step 1 (US: 236 *vs* 218, $P = 0.005$; independent: 237 *vs* 228, $P = 0.001$) and step 2 (US: 248 *vs* 232, $P = 0.006$; independent: 245 *vs* 234, $P < 0.001$) scores than unmatched applicants. There was no difference in match status when evaluating the number of applications, research experience, publications, work or volunteer experience, PhD or graduate degree. Matched US applicants were more likely to belong to AOA (17.1% *vs* 1.6%, $P = 0.002$) or attend a top 40 National Institutes of Health-funded school (31.0% *vs* 19.4%, $P = 0.002$).

Application Changes During COVID-19 Pandemic

The change to virtual interviews in 2020 resulted in an overall increase in applications submitted for all applicants. Matched 2011–2019 US applicants submitted 34–51 applications compared to a median of 56 applications in 2020 and 62 in 2021. From 2011 to 2019, matched independent applicants submitted 49–90 applications compared to a median of 84 applications in 2020 and 84–90 in 2021, consequently resulting in an increased number of interviews completed after 2020. From 2011 to 2019, matched US applicants completed 12–14 interviews compared to 16 in both 2020 and 2021. Matched independent applicants completed 4–9 interviews from 2011 to 2019 compared to 4–10 interviews in 2020 and 5–13 in 2021.

Race/Ethnicity Characteristics

Assessing 2011–021 trends, there was a significant increase in the number of applicants who identified as White (39.2%–48.3%, tau = 0.71, $P = 0.003$) or Hispanic/Latino (8.6%–11.5%, tau = 0.75, $P < 0.001$). There was no difference in the proportion of applicants identifying as Black and a decrease in those identifying as Asian (23.1%–20.9%, tau = –0.53, $P = 0.029$) or other/unknown (21.1%–8.8%, tau = –0.89, $P = 0.003$). There was no difference in the race/ethnicity proportions of active GSRs over the same period. Comparing proportions of GS applicants to residents, there was a significant drop in representation among those who identified as Asian (20.9% of applicants *vs* 12.3% of active residents, $P < 0.001$), Black/African American (8.8% of applicants *vs* 5.0% of active residents, $P < 0.001$), or Hispanic/Latino (9.4% of applicants *vs* 5.0% of active residents, $P = 0.001$) (Table 2 and Figure 1). Combining Black/African American, Hispanic/Latino, American Indian, and Native Hawaiian/Pacific Islander to form the URM category, there was a similar drop in representation of active GSRs compared with

Characteristics of Matched and Unmatched General Surgery applicants from 2011 to 2021

Characteristic	US Applicants (N = 6212)†			Independent Applicants (N = 3947)†‡		
	Matched (N = 5201)	Unmatched (N = 1016)	P Value	Matched (N = 1426)	Unmatched (N = 2521)	P Value
Applications§	48.0 (40.0, 54.0)	60.0 (52.0, 60.0)	0.177	77.0 (75.0, 84.0)	80.0 (70.0, 100.0)	0.675
Interview offers§	18.0 (17.0, 18.5)	8.0 (7.5, 12.5)	*0.005	8.0 (4.0, 9.0)	3.0 (3.0, 3.0)	*0.009
Interviews attended§	14.0 (13.0, 15.0)	8.0 (7.0, 10.0)	*0.017	7.0 (4.0, 8.0)	3.0 (3.0, 3.0)	*0.009
Programs ranked§	14.0 (12.5, 15.0)	8.0 (7, 10.5)	*0.020	6.2 (4.3, 7.8)	2.7 (2.4, 3.6)	*<0.001
Research experience	3.4 (3.1, 3.9)	2.9 (2.6, 3.3)	0.260	2.8 (2.2, 3.2)	2.7 (2.1, 2.9)	0.368
Publications‖	5.5 (4.5, 6.9)	4.0 (2.9, 4.8)	0.172	5.0 (4.1, 10.5)	5.1 (3.6, 8.6)	0.581
Work experience	3.3 (3.1, 3.5)	3.4 (3.2, 3.6)	0.810	4.3 (3.6, 4.8)	4.2 (3.9, 4.8)	0.765
Volunteer experience	6.9 (6.7, 8.0)	6.6 (6.2, 7.5)	0.240	4.9 (4.3, 5.8)	4.4 (4.0, 5.1)	0.312
AOA, %	17.1 (2.5)	1.6 (0.5)	*0.002	—	—	—
Top 40 NIH-funded school, %	31.0 (1.8)	19.4 (3.7)	*0.002	—	—	—
PhD degree	1.9 (0.3)	2.6 (0.9)	0.229	1.3 (1.8)	1.1 (1.4)	0.777
Graduate degree	18.6 (4.2)	20.3 (3.6)	0.423	23.1 (5.0)	27.0 (4.3)	0.094
USMLE score						
Step 1	236 (233, 237)	218 (214, 219)	*0.005	237 (232, 239)	228 (223, 232)	*0.001
Step 2	248 (246, 249)	232 (227, 234)	*0.006	245 (243, 246)	234 (232, 236)	*<0.001

*Denotes statistical significance with *P* < 0.05.
†N applicants who responded to the NRMP Applicant Survey.
‡Independent includes allopathic medical school graduates, US citizen and non-US citizen students and graduates of international medical schools, students and graduates of osteopathy, students and graduates of Canadian medical schools, and graduates of Fifth Pathway programs.
§Excludes Independent applicants from 2020–2021 to 2021–2022 given difference in reporting format.
‖Includes abstracts/presentations/publications.

Race/ethnicity proportions of general surgery applicants and active general surgery residents, 2011–2021

Race/Ethnicity	GS Applicants	Active GS Residents	P Value
Asian	20.9 (20.2–24.0)	12.3 (12.0–13.4)	*<0.001
URM	19.1 (18.5–19.5)	10.3 (9.6–11.4)	*<0.001
Black/African American	8.8 (7.9–9.5)	5.0 (4.5–5.1)	*<0.001
Hispanic/Latino	9.4 (8.5–10.3)	5.0 (4.9–5.8)	*0.001
White	48.3 (42.6–50.1)	46.9 (46.1–50.6)	0.478
Other/unknown†	11.0 (10.2–12.9)	30.1 (24.4–26.9)	*0.001

*Denotes statistical significance with *P* < 0.05; URM includes American Indian, Black/AA, Hispanic/Latino, Native Hawaiian/Pacific Islander.
†Other/unknown includes other, unknown, non-US citizen/nonpermanent resident, and multiple race/ethnicity (starting in 2021).

applicants (19.1% of applicants *vs* 10.3% of active residents, *P* < 0.001). Comparatively, there was a significant increase in the other/unknown group (11.0% of applicants *vs* 30.1% of active residents, *P* = 0.001) and no difference among White representation (48.3% of applicants *vs* 46.9% of active residents, *P* = 0.478).

Program Director Perspective

The distribution of factors cited by program directors for offering an interview and ranking an applicant can be visualized in Figure 2. Top cited measures for offering interviews include step 1 score (92.7%), letters of recommendation (LOR) (90.2%), and step 2 score (83.0%). Interactions with faculty during the interview/visit (92.5%), interpersonal skills (91.9%), and interactions with house staff during the interview/visit (86.8%) were the most cited factors for ranking an interviewee.

Paths Pursued by the Unmatched

During the study period, 9.9% of US and 55.0% of independent applicants did not match into a GSR position. When assessed on a 5-point Likert scale, unmatched applicants from 2013 to 2019 were most likely to participate in the Supplemental Offer and Acceptance Program (SOAP) for a GSR position (4.51/5), SOAP for a preliminary position (4.01/5), or pursue research (3.03/5) and re-apply the next cycle (4.01/5).

DISCUSSION

In a retrospective review of GSR applicants and residents, higher USMLE scores, being granted more interviews, and ranking more programs were associated with successful match outcomes for both US and independent applicants. Program directors factored step scores, LORs, and interactions/interpersonal skills when interviewing and ranking an applicant. Concerningly, there was a decrease in the proportion of active non-White GSRs compared with applicants during the 10-year study period. Excluding USMLE scores, metrics used to rank students are highly subjective and susceptible to bias; thus, programs will need to be intentionally proactive to address biases in the application and interview process.

By comprehensively assessing all components of an individual's application, a holistic approach to evaluating GSR applications is integral to diversifying the surgical workforce. Our study found that USMLE scores were significantly higher for matched than unmatched applicants. This is congruent with prior literature, as many programs still use USMLE step 1 and step 2 scores as initial screening thresholds, although there is limited data to suggest a correlation between scores and GSR success.[9,10,17,18] With the change of step 1 to pass/fail, results from a survey of program directors suggest that Step 2 may subsequently be more heavily factored.[7] However, when programs rely solely on score thresholds to screen applicants, they may have unintended consequences on interviewing and ranking a diverse pool of applicants as scores can be influenced by environmental and nonacademic factors, including race/ethnicity, socioeconomic status, and medical school.[12,19–22] Other important metrics identified in this study for interviewing an applicant include LORs, personal statements, and clerkship grades. However, each of these are known to be highly subject to bias.[23–28] In LORs,
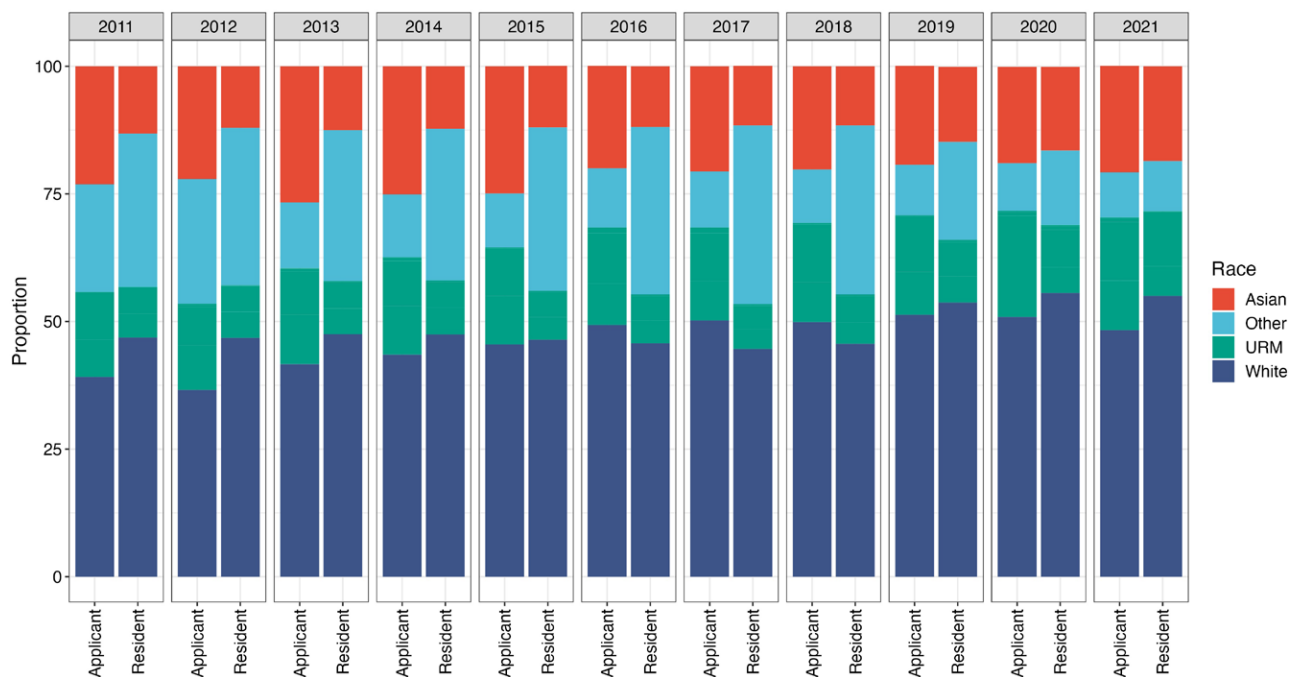
FIGURE 1. Race/ethnicity proportions of applicants and matriculants to categorical general surgery from 2011 to 2021.
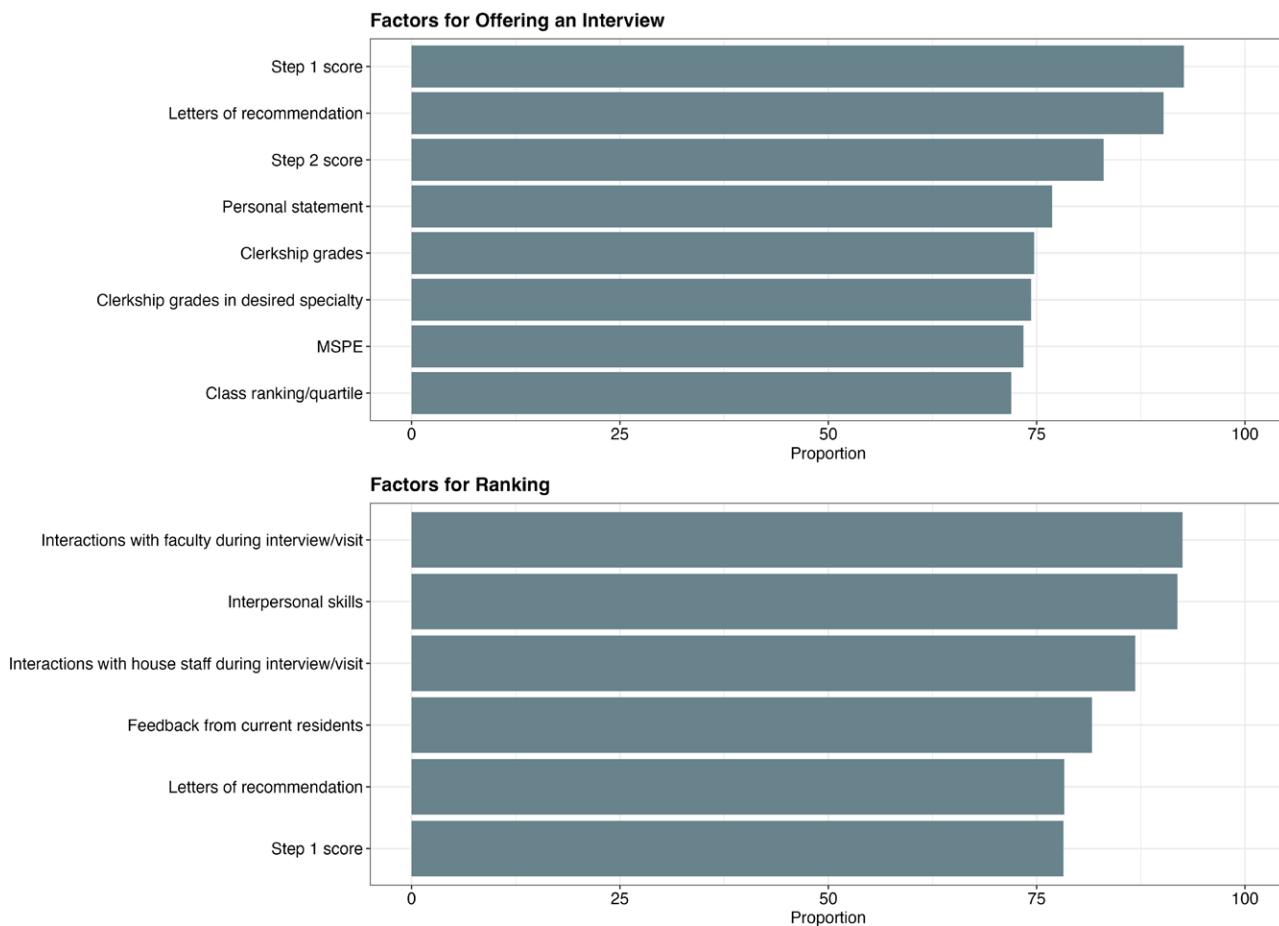

FIGURE 2. Top factors cited by program directors for (A) offering an applicant an interview and (B) ranking an applicant.

standout adjectives (eg, "excellent", "exceptional") are more likely to be used when describing White and male applicants while grindstone adjectives (eg, "dedicated", "hardworking") are more likely to be used when describing URM and female applicants.[29,30] Similarly, prior work has demonstrated a lack of inter-rater reliability when evaluating the personal statement—a

single statement could receive contradictory comments from separate reviewers.[31] Finally, clerkship grades vary widely depending on medical school, region of the country, and race/ethnicity with the percentage of students achieving honors in their surgery clerkship ranging from 6.5% to 78% across the country, favoring students from the top 20 medical schools.[24,32,33] Prior work has also suggested that non-White race and self-identified disadvantaged backgrounds are associated with lower grades in all clerkships.[12,33] Thus, programs may be inadvertently limiting the diversity of their residency class through nonstandardized evaluation of LORs, personal statements, and overemphasizing clerkship grades. This is supported by our finding of a discrepancy in URM representation among active GS residents.

Students who receive and attend more interviews are more likely to successfully match. Furthermore, interactions with faculty and house staff during the interview process were top-factored measures when ranking an applicant. However, perceptions of interpersonal interactions are subjective and heavily influenced by perceived differences in culture and background that may negatively affect candidates of differing age, gender, sex, race/ethnicity, sexual identity, religion, geographic region, family background, and physical ability.[34] During the selection process, an applicant's interpersonal skills are often referred to as an applicant's "fit" with the program. However, "fit" is highly susceptible to implicit bias, particularly against applicants who may not share the same sex, ethnicity, religion, socioeconomic status, or sexual orientation as the individual conducting the evaluation.[35,36] Prior familiarity with applicants is correlated with application review and interview scores, suggesting more interpersonal bias and underscoring the importance of standardizing the interview process to minimize potential bias.[37] It is unclear where in the application process URM students may be lost, but biases in the interview and ranking process warrant further evaluation when considering our finding of a discrepancy in active URM GS residents.

While the shift away from scored USMLE examinations is intended to promote a more holistic review of GS applicants, actualizing this vision will come with challenges. Several initiatives show promise to achieving this goal.[38,39] The switch to a virtual format eliminated the financial burden associated with in-person interviews, likely contributing to the increase in the number of applications submitted and interviews attended observed in this study. Continuation of this may promote equity in the ability for individuals to attend interviews regardless of socioeconomic status.[40,41] Additional proposed interventions to facilitate holistic review include appointing diverse interview committees, deemphasizing USMLE scores and overall grades, emphasizing important program characteristics, and standardizing interview scoring.[42] The challenge is each program receives thousands of applications for an average of 5 categorical spots every cycle, with some programs receiving over 150 applications per position.[43] Artificial intelligence may augment the holistic review of GSR applications, by leveraging large language models that can provide an initial summative evaluation of the range of data provided by applicants.[44] Furthermore, the implementation of program and regional preference signaling into the match process may be used to show applicants' strong interest in particular programs. Recommendations from the 2023 Association of Program Director in Surgery to standardize this information include implementing interview release dates to prevent interview "hoarding,"[45] providing applicants with additional information before accepting an interview so they can focus on programs they are genuinely interested in, thus allowing interviews to be more evenly distributed across all applicants. The AAMC should also consider limiting the number of applications submitted and interviews attended to minimize the burden of holistically reviewing all applicants.[11,46]

All data obtained from the NRMP and ACGME surveys are self-reported and limited to those who responded to the surveys, limiting the generalizability of the results. Current AAMC demographic data is limited to the number of students applying to each specialty and the total number of active residents in a specific specialty. Thus, our results are limited to a comparison of race/ethnicity proportions of applicants to active residents, resembling prior work.[15] Based on the available data, URM students who applied to a GSR during the study period had a significantly lower proportion of active enrollment in a GSR program. Although this does not directly correspond to a lower proportion of URM students matching into a GS program, it is imperative to assess what happens to students lost along the pipeline. There was also an increase in the proportion of active GSR self-identifying as "other" race/ethnicity. While it is possible that this may explain some of the drop in minority representation, it is unlikely that people who self-identified as minorities would change to selecting "other." Further, the drop in representation was consistent across all non-White racial groups and may remain true. Our findings also show that unmatched applicants are likely to participate in SOAP for a GSR, a preliminary position, and re-apply the next cycle or pursue research and re-apply the next cycle. However, there are no formal programs to mentor and develop students who go unmatched. Programs designed to help URM students navigate reapplying are important for supporting these students. Finally, this study only spans 2 years of data from the COVID-19 pandemic, and as such may not reflect current trends in the match. The shift toward virtual interviews will continue to have ramifications for years to come.

CONCLUSIONS

As residency programs transition to evaluating applications via pass/fail, subjective metrics (eg, LOR) may be more heavily weighted. Interactions with faculty, staff, and residents were important for ranking interviewees but may be subject to inherent bias; this is a cause for concern to diversify the surgical workforce. Structural barriers affecting match status warrant further evaluation. Future studies may focus on what characteristics predict success when applicants matriculate as GSRs.

REFERENCES

1. Association of American Medical Colleges. Advancing Diversity, Equity, and Inclusion in Medical Education, 2023. Available at: https://www.aamc.org/about-us/equity-diversity-inclusion/advancing-diversity-equity-and-inclusion-medical-education. Accessed June 12, 2023.
2. Cohen JJ, Gabriel BA, Terrell C. The case for diversity in the health care workforce. *Health Aff (Millwood)*. 2002;21:90–102.
3. Marrast LM, Zallman L, Woolhandler S, et al. Minority physicians' role in the care of underserved patients: diversifying the physician workforce may be key in addressing health disparities. *JAMA Intern Med*. 2014;174:289–291.
4. General Surgeon Demographics and Statistics in the US, 2023. Available at: https://www.zippia.com/general-surgeon-jobs/demographics/. Accessed March 2, 2023.
5. United States Census Bureau. United States Census Bureau Quick Facts, 2023. Available at: https://www.census.gov/quickfacts/fact/table/US/PST045222. Accessed May 1, 2023.
6. Stain SC, Hiatt JR, Ata A, et al. Characteristics of highly ranked applicants to general surgery residency programs. *JAMA Surg*. 2013;148:413–417.
7. Aziz H, Khan S, Rocque B, et al. Selecting the next generation of surgeons: general surgery program directors and coordinators perspective on USMLE changes and holistic approach. *World J Surg*. 2021;45:3258–3265.
8. Dudas LM, Khan US, Borgstrom DC. Mastering the match: recruiting a successful surgery resident. *Curr Surg Rep*. 2021;9:27.
9. Iwai Y, Lenze NR, Becnel CM, et al. Evaluation of predictors for successful residency match in general surgery. *J Surg Educ*. 2022;79:579–586.
10. 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;11:268–273.
11. Naaseh A, de Virgilio C, Nahmias J. General surgery residency match: time for more than a virtual change. *J Surg Educ*. 2021;78:1771–1775.

12. Henderson MC, Jerant A, Unkart J, et al. The relationships among self-designated disadvantage, socioeconomic disadvantage, and academic performance in medical school: a multi-institutional study. *J Health Care Poor Underserved*. 2020;31:208–222.

13. Jerant A, Sciolla AF, Henderson MC, et al. Medical student socioeconomic disadvantage, self-designated disadvantage, and subsequent academic performance. *J Health Care Poor Underserved*. 2019;30:1419–1432.

14. Williams M, Kim EJ, Pappas K, et al. The impact of United States Medical Licensing Exam (USMLE) step 1 cutoff scores on recruitment of underrepresented minorities in medicine: a retrospective cross-sectional study. *Health Sci Rep*. 2020;3:e2161.

15. Caruso JP, Singh R, Mazzola C, et al. Trends in racial and ethnic representation among neurosurgery applicants and residents: a comparative analysis of procedural specialties. *World Neurosurg*. 2022;163:e177–e186.

16. Collins RA, Sheriff SA, Yoon C, et al. Assessing the complex general surgical oncology pipeline: trends in race and ethnicity among US medical students, general surgery residents, and complex general surgical oncology trainees. *Ann Surg Oncol*. 2023;30:4579–4586.

17. Shirkhodaie C, Avila S, Seidel H, et al. The association between USMLE step 2 clinical knowledge scores and residency performance: a systematic review and meta-analysis. *Acad Med*. 2023;98:264–273.

18. Sajadi-Ernazarova K, Ramoska EA, Saks MA. USMLE scores do not predict the clinical performance of emergency medicine residents. *Mediterr J Emerg Med Acute Care*. 2020;1:4–7.

19. Rubright JD, Jodoin M, Barone MA. Examining demographics, prior academic performance, and United States medical licensing examination scores. *Acad Med*. 2019;94:364–370.

20. Neville AL, Smith BR, de Virgilio C. USMLE step 1 scoring system change to pass/fail—an opportunity for change. *JAMA Surg*. 2020;155:1093–1094.

21. Tatem GB, Gardner-Gray J, Standifer B, et al. While you don't see color, I see bias: identifying barriers in access to graduate medical education training. *ATS Sch*. 2021;2:544–555.

22. Quesada PR, Solis RN, Ojeaga M, et al. Overemphasis of USMLE and its potential impact on diversity in otolaryngology. *OTO Open*. 2021;5:2473974X211031470.

23. Teherani A, Hauer KE, Fernandez A, et al. How small differences in assessed clinical performance amplify to large differences in grades and awards: a cascade with serious consequences for students underrepresented in medicine. *Acad Med*. 2018;93:1286–1292.

24. Low D, Pollack SW, Liao ZC, et al. Racial/ethnic disparities in clinical grading in medical school. *Teach Learn Med*. 2019;31:487–496.

25. Rojek AE, Khanna R, Yim JW, et al. Differences in narrative language in evaluations of medical students by gender and under-represented minority status. *J Gen Intern Med*. 2019;34:684–691.

26. Ross DA, Boatright D, Nunez-Smith M, et al. Differences in words used to describe racial and gender groups in Medical Student Performance Evaluations. *PLoS One*. 2017;12:e0181659.

27. Drake C, Lewis CF, Lekas H-M. Reckoning with racism in the match process. *Acad Psychiatry*. 2022;46:663–666.

28. Girgis MY, Qazi S, Patel A, et al. Gender and racial bias in letters of recommendation for orthopedic surgery residency positions. *J Surg Educ*. 2023;80:127–134.

29. Turrentine FE, Dreisbach CN, St Ivany AR, et al. Influence of gender on surgical residency applicants' recommendation letters. *J Am Coll Surg*. 2019;228:356–365.e3.

30. Powers A, Gerull KM, Rothman R, et al. Race- and gender-based differences in descriptions of applicants in the letters of recommendation for orthopaedic surgery residency. *JB JS Open Access*. 2020;5:e20.00023 –e20.00023.

31. White BAA, Sadoski M, Thomas S, et al. Is the evaluation of the personal statement a reliable component of the general surgery residency application? *J Surg Educ*. 2012;69:340–343.

32. Quinn KM, Campbell L, Mukherjee R, et al. Step 1 is pass/fail, now what? Can clinical clerkship grades be used as a reliable metric to screen general surgery residency applicants? *J Surg Res*. 2022;279: 592–597.

33. Lee KB, Vaishnavi SN, Lau SK, et al. "Making the grade:" noncognitive predictors of medical students' clinical clerkship grades. *J Natl Med Assoc*. 2007;99:1138–1150.

34. DeVito JA. *Culture and Interpersonal Communication The Interpersonal Communication Book*. SAGE Publications, Inc, 2022.

35. Phillips MR, Charles A. Addressing implicit bias in the surgical residency application and interview process for underrepresented minorities. *Surgery*. 2021;169:1283–1284.

36. Chapman EN, Kaatz A, Carnes M. Physicians and implicit bias: how doctors may unwittingly perpetuate health care disparities. *J Gen Intern Med*. 2013;28:1504–1510.

37. Towaij C, Gawad N, Alibhai K, et al. Trust me, I know them: assessing interpersonal bias in surgery residency interviews. *J Grad Med Educ*. 2022;14:289–294.

38. Parker AS, Mwachiro MM, Kirui JR, et al. A semistructured interview for surgical residency targeting nontechnical skills. *J Surg Educ*. 2022;79:e213–e219.

39. Dossett LA, Mulholland MW, Newman EA; Michigan Promise Working Group for Faculty Life Research. Building high-performing teams in academic surgery: the opportunities and challenges of inclusive recruitment strategies. *Acad Med*. 2019;94:1142–1145.

40. Sarac BA, Janis JE. Perspectives of virtual residency interviews in plastic surgery: results following 1 year of training. *Plast Reconstr Surg Glob Open*. 2023;11:e4746.

41. Dey P, Havlik JL, Kurup V, et al. Lessons from the COVID-19 pandemic: the case for virtual residency interviews. *Ann Surg Open*. 2021; 2:e077.

42. Nehemiah A, Roberts SE, Song Y, et al. Looking beyond the numbers: increasing diversity and inclusion through holistic review in general surgery recruitment. *J Surg Educ*. 2021;78:763–769.

43. Weissbart SJ, Kim SJ, Feinn RS, et al. Relationship between the number of residency applications and the yearly match rate: time to start thinking about an application limit? *J Grad Med Educ*. 2015;7:81–85.

44. Schild I. What is ChatGPT and how are employers are to using it in Recruiting? Foundation List Nonprofit Jobs - Nonprofit, Foundation and Education Job Board; 2023. Available at: https://www.foundation-list.org/chatgpt-and-recruiting/. Accessed July 10, 2023.

45. McKinley SK, Altieri MS, Sheppard O, et al; ASE Graduate Surgical Education Committee. Designing the "match of the future": challenges and proposed solutions in the interview and match phase of the UME–GME transition. *Global Surg Educ*. 2022;1:69.

46. Whipple ME, Law AB, Bly RA. A computer simulation model to analyze the application process for competitive residency programs. *J Grad Med Educ*. 2019;11:30–35.