

Original Research

Incorporating Senior Residents in a Holistic General Surgery Applicant Review Process

Andrew Mudreac, M.D.¹, Solange Bayard, M.D.², Genevieve A. Fasano, M.D.¹, Josh Johnson, M.D.¹, Alexander Kossar, M.D.³, Alexey Abramov, M.D.⁴, Casidhe-Nicole Ramirez Bethancourt, M.D.⁵, Stephen Oh, M.D.¹, Beth Hochman, M.D.⁶

¹ Department of Surgery, Weill Cornell Medicine, ² Department of Surgery, Memorial Sloan Kettering Cancer Center, ³ Division of Cardiothoracic Surgery, Emory University Hospital, ⁴ Divisions of Cardiac and Thoracic Surgery, UCLA Health, ⁵ Department of Surgery, Columbia University, ⁶ Department of Surgery, New York University

Keywords: general surgery, implicit bias, internship and residency, medical education, medical students, selection bias

International Journal of Surgical Education

Background

We aimed to restructure the application review process to design a standardized and holistic review process that includes senior residents in evaluating a broader pool of applicants for a general surgery residency program.

Methods

In 2020 and 2021, a new quantitative reviewer scoresheet was created and optimized for objectively evaluating general surgery residency applicants based on the core values of the program. Faculty and senior residents were invited to review applications and conduct interviews. The quality and effectiveness of the selection process was assessed by analyzing scoring rubrics, reviewer feedback, and match results.

Results

Most reviewers reported that the scoresheet was easy to use (90%), was a fair approach for application review (95%), and accurately reflected the qualities needed for a surgical resident to be successful in the program (74%). All reviewers with prior experience in application review reported that the new process was at least equivalent to the prior approach. Mean faculty and resident scores were not significantly different in 2020 (faculty mean, 7.21; resident mean, 7.2; $P = .78$) or 2021 (faculty mean, 6.72; resident mean, 7.17; $P = .11$). In both years, most (82% to 85%) top scorers remained in the top tier of our final rank list.

Conclusions

We established a standardized, holistic process for reviewing general surgery applicants that includes senior surgical residents. This standardized process can be used as a model for other residency programs that wish to develop a standardized process for selecting residents to match to their program.

INTRODUCTION

For the past 70 years, residency programs have selected prospective residents through the National Resident Matching Program.¹ For residency programs, the process of evaluating and ranking residency applicants has traditionally relied on medical school reputations and numerical data, such as standardized test scores and grades. Although scores and grades are theoretically objective and can quickly screen thousands of applications, these factors have not strongly predicted clinical performance. They are also subject to structural racism and unconscious bias.²⁻⁶

In 2022, the United States Medical Licensing Examination (USMLE) switched its Step 1 exam from a numeric metric commonly used to filter applications to a pass/fail out-

come.^{6,7} However, the USMLE Step 2 Clinical Knowledge (CK) exam still uses a numerical score. As a result, this score may be evaluated as a metric in an initial application screen⁸⁻¹¹ and, thus, may contribute to the same pitfalls as the historical Step 1 exam.

In 2020 and 2021, our general surgery program, located at a large tertiary care hospital in the New York Metropolitan Area, received more than 1400 applications each year to fill 7 categorical and 5 preliminary positions. This volume of applications has historically been untenable for faculty to sift through without first using automatic numerical filters and is difficult for a small group to review within a tight timeframe. Even after culling the applicant pool, we did not have a standard approach to holistically review all parts of an application, nor did we have a way to evaluate comparable, consistent, or trackable measures among reviewers.

Without a standard approach and more resources, the review process would remain complicated, consume time, and risk perpetuating bias.

One strategy to expand review resources is to include residents. Residents have the capacity to help with the selection process and are a reliable resource for evaluating applications.¹² They can also augment the pool of reviewers, reduce the individual reviewer workload, expand the diversity of perspective when considering applicants, and enhance engagement with their own experiences. Also, when residents are included in the review process, programs acknowledge their vital role in defining the landscape of a residency program.

In this study, we aimed to restructure the application review process to create a simplified, holistic approach that incorporates residents in evaluating a broader pool of residency applicants. Our goals were to develop a review process that is (1) fair in supporting an equitable and inclusive evaluation of applicants that limits bias; (2) feasible and easy to use; (3) accurate in assessing what we think are important qualities in trainees; (4) preferable or equal to the prior process; (5) precise in generating transparent, objective, and replicable scores with acceptable variability between scorers; and (6) non-inferior in matching from the top tier of our rank list.

METHODS

SCORING RUBRIC DEVELOPMENT

We convened a committee of faculty and residents to determine selection criteria for applicants to our general surgery residency program. These criteria were based on the goals and values of the residency program, the mission of the hospital, and the core competencies of the Accreditation Council for Graduate Medical Education (ACGME). The committee agreed that ideal applicants embodied 5 core categories of equal weight: service/teaching, education/life-long learning, leadership/teamwork, research/scholarly activity, and perseverance/commitment.

With these core categories, we developed a standardized rubric for scoring Electronic Residency Application Service (ERAS) applications and interviews. The rubric was optimized each year until the final rubric was approved in 2022 (Fig. 1). The improvements included adjusting the appearance and approach to the embedded calculations based on feedback from applicant reviewers.

On the rubric, each category was quantified on a 5-point Likert scale from “Red Flag” to “Outstanding”. Each category was also clarified with a guiding question and possible examples of candidate (1) activities and accomplishments that can be gleaned from the application and (2) behaviors that can be noted in interviews. All categories were assigned equal weight. To be fully transparent with participants, we also included our filters, reviewer composition, and scoring implications at the top of the rubric. We included an optional field “Should we interview?” to cross-check that our rubric generated commensurately high scores for applicants whom reviewers wanted to interview.

To improve our final ranking strategy, we also included the optional field: “How should we rank?” (Fig. 1).

APPLICATION REVIEWS AND INTERVIEWS

To fully represent our program, we invited both faculty and senior residents to review applications and conduct interviews. All core educational faculty were required to participate. This greater volume of reviewers ensured we could review an expanded pool of applications and that each reviewer could manage the workload (30-50 applications) despite their other clinical, academic, and administrative duties. To recognize applicants’ and reviewers’ diversity of background and experience, we assigned at least 2 reviewers to each application. To maximize the number of perspectives, applicants invited for interviews met with 3 interviewers who were different from their application reviewers. Each application had no more than 1 resident reviewer, and each applicant being interviewed had no more than 1 resident interviewer. All participants completed training on how to use the scoring rubrics. However, we were not too prescriptive with scoring to respect and embrace the different views of interviewers, and to support ranking a diverse pool of residents.

After application reviews, we consolidated scoresheets to generate a composite master score for each applicant. We predetermined that interviews would be granted to all top scorers in addition to all internal applicants and subinterns to fill a set number of interview slots. After interviews, sum scores were again calculated and averaged between interviewers for each applicant, which were then used to create a semifinal rank list. Program leadership, faculty, and senior residents then jointly discussed all data points and adjusted the list to create a final rank list for matching.

BIAS MITIGATION

To minimize explicit and implicit biases among reviewers and interviewers, we took several measures. Using ERAS filters, we selected all candidates who passed the USMLE Step 1 rather than setting a specific score cutoff. Also, we included any candidates who graduated medical school within the past 2 years rather than focusing only on applicants who were currently in their last year of medical school. To purposefully structure the application content seen by reviewers, we sent packets to reviewers as PDFs rather than giving reviewers direct access to ERAS. Although we could not remove the applicant names or pronouns from the packets, we omitted applicant photos, placed *curriculum vitae* and personal statements at the beginning of each PDF, and moved medical school grades and USMLE exam scores to the end of each PDF. To prevent bias from seeing others’ scores and comments, we gave each reviewer separate scoresheets and discouraged them from discussing applicants before submitting their scoresheets for data compilation.

To encourage standardized interviews, we suggested value-based interview questions. Also, all interviewers completed training on implicit bias developed by our graduate medical education office. In 2020, we considered grant-

2022 APPLICATION REVIEW								
Categories informed by academic mission, ACGME competencies, and our own self-defined program values **everyone interprets the meanings on their own when reviewing because everyone has their own internal consistency. LORs and MSPE may inform the scores chosen; they are not separate categories** INVITED REVIEWERS: all faculty (PEC and CCC required), PGY5s, PGY4s, lab residents FILTERS: passed USMLE1, J1/O visa or otherwise able to work in US, graduated medical school within 2 years of submitting application Other considerations: prelims and VPS students								
BELOW ARE THE CATEGORIES BEING SCORED, APPLICATION EXAMPLES, and ASSOCIATED VALUES								
REVIEWER GROUP	Service/ Teaching	Education/ Life-long Learning	Leadership/ Teamwork	Research/ Scholarly Activity	Perseverance/ Commitment	SCORE (0-10)	SHOULD WE INTERVIEW?	COMMENTS
Last, First Name (pink = prelim applicant)	CAN YOU TEACH OTHERS? (eg volunteer work, public service, mentoring, tutoring or teaching, serving as a caregiver) ACGME: patient care Program: applied health care, clinical skills Value: inclusion, self and situational awareness	CAN YOU BE TAUGHT? (eg pursuit of higher degrees, tier of school/quantile of class, scholarship recipient, learning a new skill as an adult) ACGME: medical knowledge, practice-based learning Program: education Value: judgment, curiosity	CAN YOU LEAD? (eg chair of a committee, involvement in team activity including sports, role as manager or leader, ability to make fun of oneself / have insight into weaknesses) ACGME: interpersonal and communication skills Program: empathy, cultural diversity, communication, understanding one's own limits Value: humility, honesty, communication	CAN YOU PROBLEM-SOLVE? (eg abstracts, presentations, publications, innovations, patents, products for public consumption including arts/music) ACGME: systems-based practice Program: research, adaptability Value: adaptability, innovation	ARE YOU RELIABLE? (eg overcoming an obstacle, going above and beyond, taking a task to absolute completion, sticking with a particular activity for a long time, working or supporting dependents while in school, starting a company, running a marathon) ACGME: professionalism Program: resilience Value: commitment, confidence	>9 Interview likely <7.5 Interview unlikely	(c)	
	FOR EACH CATEGORY, TYPE AN INTEGER SCORE FROM 0-4 AS FOLLOWS Outstanding = 4 Above average = 3 Average = 2 Below average = 1 Red flag = 0						DO NOT TOUCH this is auto-calculated	CHOOSE FROM DROPDOWN Yes Maybe No
Group 1	0	1	2	3	4	5	(b)	

2022 INTERVIEW SCORESHEET								
Categories informed by academic mission, ACGME competencies, and our own self-defined program values **everyone interprets the meanings on their own when reviewing because everyone has their own internal consistency. both the application and the interview should inform the scores chosen** INVITED INTERVIEWERS: all faculty (PEC and CCC required), PGY5s, PGY4s, lab residents								
BELOW ARE THE CATEGORIES BEING SCORED, INTERVIEW/APPLICATION EXAMPLES, and ASSOCIATED VALUES								
INTERVIEW DATE	Service/ Teaching	Education/ Life-long Learning	Leadership/ Teamwork	Research/ Scholarly Activity	Perseverance/ Commitment	SCORE (0-10)	HOW SHOULD WE RANK?	COMMENTS
Last, First Name (pink = prelim applicant)	CAN YOU TEACH OTHERS? INTERVIEW: can explain thoughts clearly APPLICATION: volunteer work, public service, mentoring, tutoring or teaching, serving as a caregiver ACGME: patient care Program: applied health care, clinical skills Value: inclusion, self and situational awareness	CAN YOU BE TAUGHT? INTERVIEW: asks informed questions APPLICATION: pursuit of higher degrees, tier of school/quantile of class, scholarship recipient, learning a new skill as an adult ACGME: medical knowledge, practice-based learning Program: education Value: judgment, curiosity	CAN YOU LEAD? INTERVIEW: personable, humble, honest, uses "we" instead of "I," can recognize and enumerate others' perspectives APPLICATION: chair of a committee, involvement in team activity including sports, role as manager or leader, ability to make fun of oneself / have insight into weaknesses ACGME: interpersonal and communication skills Program: empathy, cultural diversity, communication, understanding one's own limits Value: humility, honesty, communication	CAN YOU PROBLEM-SOLVE? INTERVIEW: curious, interested in change APPLICATION: abstracts, presentations, publications, innovations, patents, products for public consumption including arts/music ACGME: systems-based practice Program: research, adaptability Value: adaptability, innovation	ARE YOU RELIABLE? INTERVIEW: driven, passionate APPLICATION: overcoming an obstacle, going above and beyond, taking a task to absolute completion, sticking with a particular activity for a long time, working or supporting dependents while in school, starting a company, running a marathon ACGME: professionalism Program: resilience Value: commitment, confidence	>=9 top tier likely <8 top tier unlikely	(d)	
	FOR EACH CATEGORY, TYPE AN INTEGER SCORE FROM 0-4 AS FOLLOWS Outstanding = 4 Above average = 3 Average = 2 Below average = 1 Red flag = 0						DO NOT TOUCH this is auto-calculated	CHOOSE FROM DROPDOWN Top Tier 2nd Tier 3rd Tier Do Not Rank
Nov	0	1	2	3	4	5	(b)	

Figure 1. Final application review and interview rubrics in 2022. The general concept was described at the top of the rubrics and examples were included under each scoring category. Important features of the final rubric included: (a) guiding questions and examples in the category fields, (b) clarifying details on the impact of scores, (c) an optional “Should we interview?” field, and (d) an optional “How should we rank?” field

CCC = Clinical Competency Committee. PEC = Program Evaluation Committee

ing interviews to applicants whose scores fell below the top tier but for whom current or former faculty provided pointedly high recommendations. Given inherent bias in this loophole, we eliminated the option altogether in our second year.

QUALITY ANALYSIS

To assess the quality of the selection process, we analyzed scoring rubrics for accuracy, precision, and variability between reviewers. We also assessed the feasibility and fairness of the review process by developing and distributing anonymous electronic surveys to reviewers and interviewers. Each year, we refined the scoring rubric based on our experience and reviewer feedback, which led to our final rubric in 2022 (Fig. 1). We also evaluated the match process by comparing numerical rubric scores to the final rank list and how highly we matched candidates within the rank list.

This study met the criteria for Institutional Review Board exemption.

THE TEAM

Project leadership comprised the director and associate director of the general surgery residency program. These leaders developed the concept and implementation strategy by engaging members of the Program Evaluation Committee (PEC), including core educational faculty and resident representatives. The associate program director created the scoring tool and instructions for users, iteratively improved the tool with input from reviewers, collated data, performed data analyses, and shared results with participants. The application reviewers and interviewers included all members of the PEC and Clinical Competency Committee (CCC), as well as other faculty and senior residents participating in the selection process.

Table 1. Reviewer survey feedback from 2020 to 2021

	Yes, No. (%)	Neutral, No. (%)	No, No. (%)	Total
Fair	40 (95)	1 (2.4)	1 (2.4)	42
Feasible (easy to use)	38 (90)	1 (2.4)	3 (7.1)	42
Preferable to prior process	12 (48)	9 (36)	4 (16)	25
Accurate (supports complete evaluation)	37 (88)	2 (4.8)	3 (7.1)	42
Accurate (measures what is considered important)	31 (74)	9 (21)	2 (4.8)	42

STATISTICAL ANALYSIS

To assess differences between mean faculty and resident scores, a Welch two-sample t-test was performed using R statistical software (version 4.3.1).

RESULTS

To learn how reviewers and interviewers felt about the standardized process, we first surveyed reviewers after each application review cycle. In 2020, 12 of 21 (57%) reviewers responded to the survey; in 2021, 21 of 40 (53%) reviewers responded. In 2021, we also surveyed interviewers after the first interview day, to which 9 of 21 (43%) interviewers responded. Most of the 42 total respondents reported that the approach was fair (40 fair, 1 neutral, 1 unfair), easy to use (38 easy, 1 neutral, 3 difficult), allowed full evaluation of applicants (37 satisfied, 2 neutral, 3 dissatisfied), and accurately reflected the qualities needed for a successful resident in our program (31 yes, 9 neutral, 2 no) (Table 1). The 25 respondents who were involved in the prior application review process indicated that the standardized process was preferable or at least equal to the prior process (12 yes, 9 neutral, 4 no). When asked why the standardized process was preferable, respondents cited objectivity, organization, fairness, and categorization.

We assessed precision by evaluating variability between completed scoresheets. In 2020, reviewers' scores were highly precise, with 91% of applications scoring within 2 points between reviewers (Fig. 2). In 2021, this precision was lower at 78%. It is pertinent to note that in 2020, the score calculation generated values from 2 to 10. In 2021, the score calculation was adjusted to generate values from 0 to 10. When reviewer scores were compared to other faculty and residents, the score distributions did not greatly differ between reviewers or among the whole group (Fig. 3, 4). Also, resident scores were similar to faculty scores and consistent with the whole group (Fig. 3, 4). Mean faculty and resident scores were not significantly different in 2020 [faculty mean, 7.21 (n = 13); resident mean, 7.2 (n = 9); $P = .78$] or 2021 [faculty mean, 6.72 (n = 27); resident mean, 7.17 (n = 13); $P = .11$].

Ultimately, we invited applicants with the highest composite scores for interviews (Fig. 5; n = 68 categorical applicants, n = 9 preliminary applicants). These composite scores did not closely correlate with the USMLE scores. We also invited applicants who met predetermined criteria

(internal applicants and subinterns) and, in 2020 only, we granted interviews to a subset of applicants whose scores fell below the top tier but for whom current or former faculty provided pointedly high recommendations. These invitations led to about 100 interviews each year (2020, n = 102; 2021, n = 105).

After the interviews, we created our preliminary and final rank lists. Most top scorers (2020, 82%; 2021, 85%) remained in the top tier of the final rank list. In both years, all categorical residents matched from within the top tier of the final rank list, as in prior years.

DISCUSSION

In this study, we standardized the application review and interview process for applicants seeking to match to our general surgery residency program. We designed this process to include valuable resident participation, mitigate bias, and focus on our program's core values and strengths. Our process was feasible to implement, easy to use, at least as good as our prior approach, and accurate in selecting the top tier of applicants for matching. These findings support that our standardized process that incorporates residents is a valuable, holistic approach for fairly, easily, and accurately selecting residency applicants for our general surgery program.

Medical schools rely on a holistic review process to recruit future physicians that align with the school's mission, reduce bias in applicant selection, and, ultimately, diversify the physician workforce.¹³⁻¹⁵ Residency programs may similarly benefit from a holistic selection process. The transition to pass/fail scoring on the USMLE Step 1 exam may have served as a catalyst to prompt residency programs to reevaluate their selection process and restructure it to fit the needs of a dynamic educational landscape. One method to restructure a program would be to simply use the remaining available application materials (eg, Step 2 CK score, number of publications, Alpha Omega Alpha status, letters of recommendation) for screening and selecting applicants to interview and rank.¹⁶⁻¹⁹ However, the USMLE score change is also an opportunity for programs to introduce a standardized, holistic review process. With such a process, residency programs can maintain the integrity of selecting highly competitive applicants with desirable, pre-defined attributes while recruiting a more diverse resident class.^{20,21}

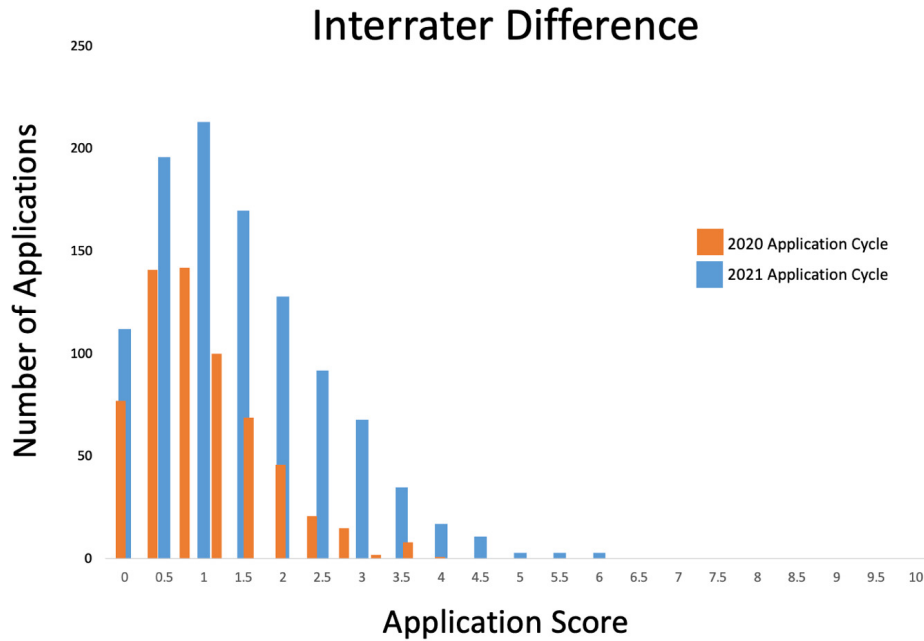
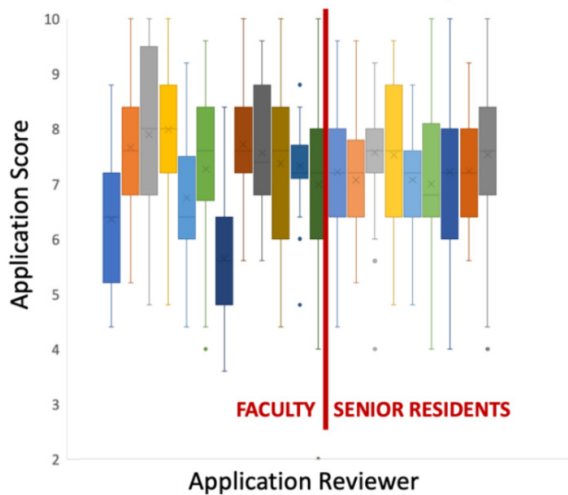


Figure 2. Distribution of differences between scores for each application. In 2020, scores could range from 2 to 10, with a maximum possible difference of 8. In 2021, scores could range from 0 to 10, with a maximum possible difference of 10

Individual Reviewer Score Ranges 2020



Individual Reviewer Score Ranges 2021

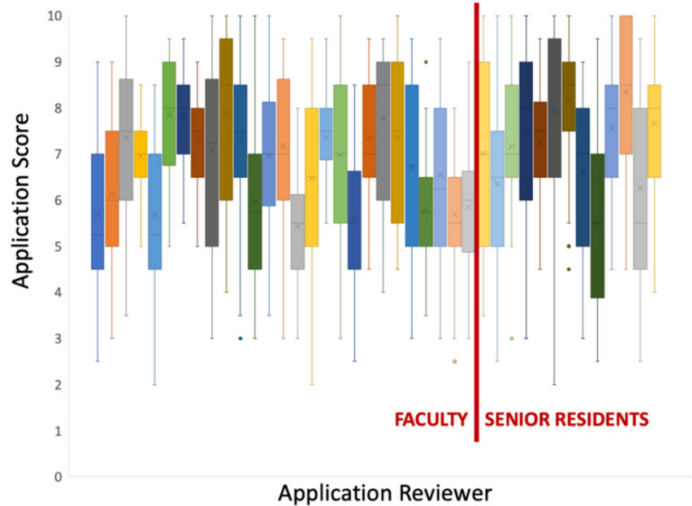


Figure 3. Box and whisker plot of score ranges generated by faculty and residents in 2020 and 2021. Each bar represents the 25th to 75th percentile ranges of a given reviewer’s scores. Each “whisker” (line extending above and below each bar) represents the bottom and top quartiles of a given reviewer’s scores. The “x” within each bar marks that reviewer’s mean, and the horizontal line within each bar represents that reviewer’s median. Dots are outlier data points

Although our review process included faculty and residents, we noticed challenges in getting these individuals to engage in educational processes. In our experience, only faculty and a select few senior residents are involved in the applicant evaluation and selection process. Although

we need to expand and diversify our reviewer pool, our new standardized approach established an easy, feasible method that will help to recruit more faculty and residents to participate in reviewing residency applications.

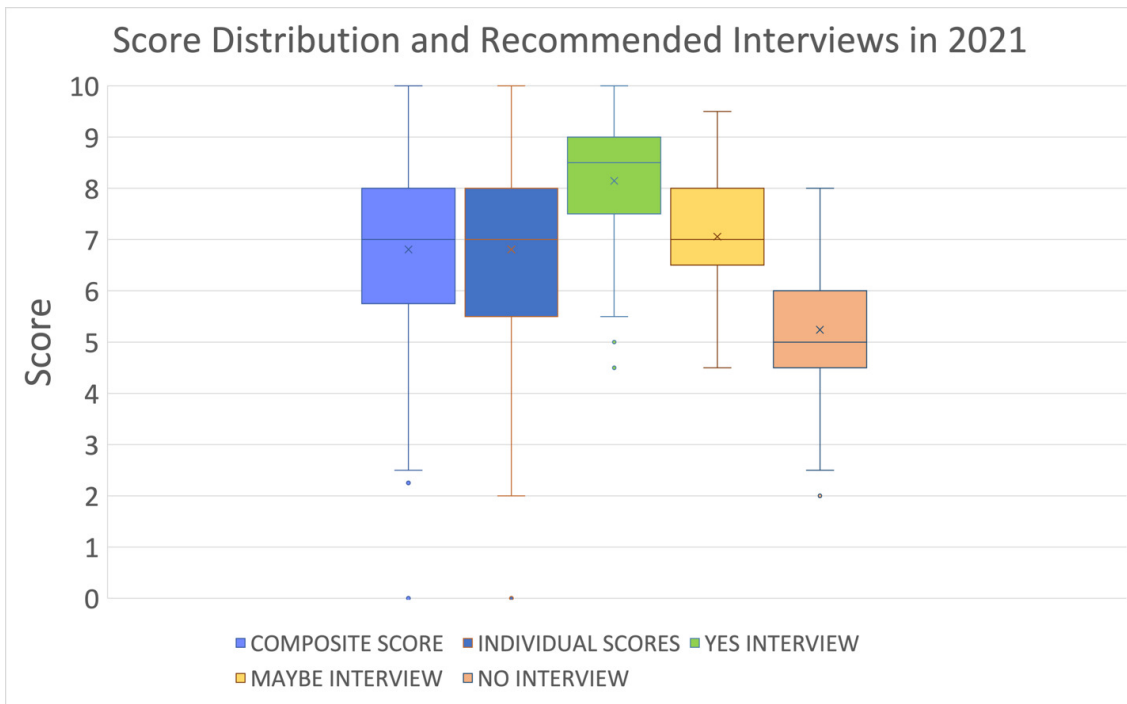


Figure 4. Individual reviewer and composite distribution of score ranges in 2021. The optional field “Should we interview?” was added in 2021 to cross-check that the rubric generated commensurately high scores for applicants whom reviewers felt strongly should be interviewed

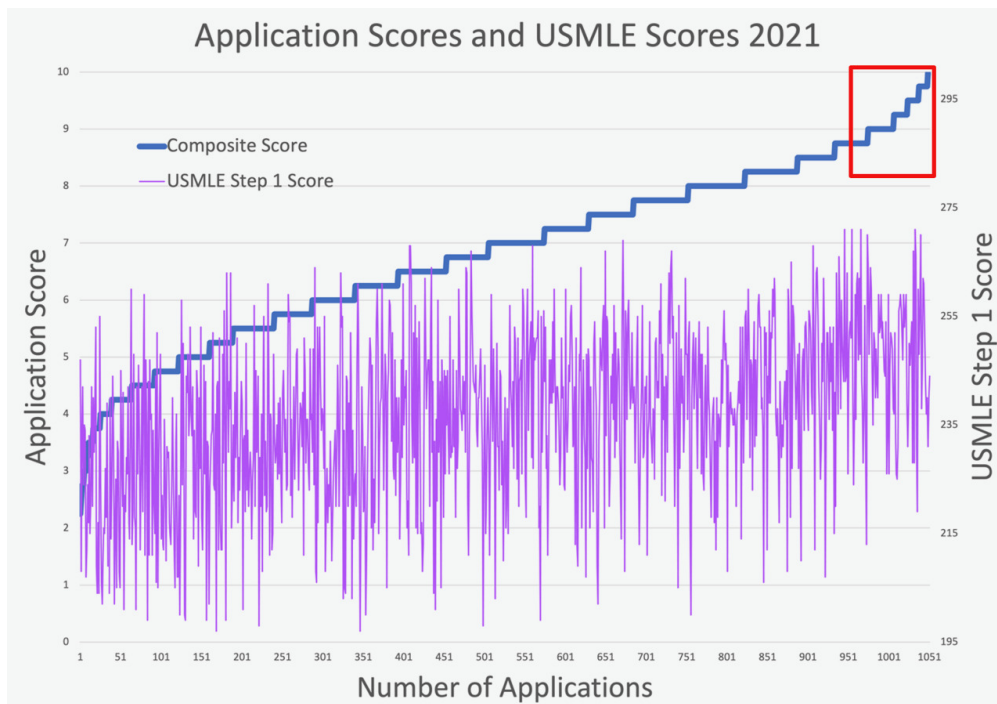


Figure 5. Distribution of application review scores in 2021. Each application had 2 reviewers whose scores were averaged to produce a final composite score. The red box represents the top scorers who were invited for interviews (68 categorical and 9 preliminary applicants)

With regard to future efforts for residency improvement, increasing overall engagement from faculty and residents in the educational process remains a challenge. Commonly, only faculty and select few senior residents are involved in the applicant evaluation and selection process. Inclusion

of a diverse group of senior residents and faculty in this process should be examined as a potential means of addressing this challenge.

Our approach exemplifies a strategy for other residency programs to develop standardized processes for selecting

residency applicants in their departments. Based on our experience in developing this process, we suggest that other residency programs consider the following recommendations. Before implementation, programs would benefit from engaging and aligning department faculty/leadership, creating a hierarchy of values, calculating department capacity for accommodating interviews, assembling a team of core reviewers and interviewers, distributing new score rubrics, educating reviewers, and designing metrics to evaluate quality improvement. After implementation, programs would benefit from taking steps to ensure the quality of the new process, assessing outlier scores and missing values, collecting feedback from reviewers, and using feedback to adjust the process.

CONCLUSIONS

We established a standardized, holistic process for reviewing applications, interviewing applicants, and accurately ranking applicants for matching in our general surgery residency program. This standardized process includes residents as valuable resources in the review process and acknowledges their important role in the educational landscape of our residency program. This successful standardized process is a valuable approach that other residency programs can use as a model for developing their own standardized process for selecting residents to match to their program.

Submitted: October 16, 2024 EST, Accepted: November 03, 2024 EST

REFERENCES

1. Roth AE. The Origins, History, and Design of the Resident Match. *JAMA*. 2003;289(7):909. [doi:10.1001/jama.289.7.909](https://doi.org/10.1001/jama.289.7.909)
2. Panda N, Bahdila D, Abdullah A, Ghosh AJ, Lee SY, Feldman WB. Association Between USMLE Step 1 Scores and In-Training Examination Performance: A Meta-Analysis. *Acad Med*. 2021;96(12):1742-1754. [doi:10.1097/ACM.0000000000004227](https://doi.org/10.1097/ACM.0000000000004227)
3. Solomon SR, Atalay AJ, Osman NY. Diversity Is Not Enough: Advancing a Framework for Antiracism in Medical Education. *Acad Med*. 2021;96(11):1513-1517. [doi:10.1097/ACM.0000000000004251](https://doi.org/10.1097/ACM.0000000000004251)
4. Cofer KD, Hollis RH, Goss L, Morris MS, Porterfield JR, Chu DI. Burnout is Associated With Emotional Intelligence but not Traditional Job Performance Measurements in Surgical Residents. *J Surg Educ*. 2018;75(5):1171-1179. [doi:10.1016/j.jsurg.2018.01.021](https://doi.org/10.1016/j.jsurg.2018.01.021)
5. Ray JJ et al. Association Between American Board of Surgery In-Training Examination Scores and Resident Performance. *JAMA Surg*. 2016;151(1):26. [doi:10.1001/jamasurg.2015.3088](https://doi.org/10.1001/jamasurg.2015.3088)
6. Belovich AN et al. USMLE Step-1 is Going to Pass/Fail, Now What Do We Do? *Med Sci Educ*. 2021;31(4):1551-1556. [doi:10.1007/s40670-021-01337-4](https://doi.org/10.1007/s40670-021-01337-4)
7. Lauer C, Buonpane C, Lane S, Foreman Z, Shabahang M. General Surgery Residency Application Evaluation in a USMLE Step 1 Pass/Fail World: A Retrospective Comparison. *J Surg Res*. 2021;265:317-322. [doi:10.1016/j.jss.2021.03.056](https://doi.org/10.1016/j.jss.2021.03.056)
8. Khalil S et al. The importance of USMLE step 2 on the screening and selection of applicants for general surgery residency positions. *Heliyon*. 2023;9(7):e17486. [doi:10.1016/j.heliyon.2023.e17486](https://doi.org/10.1016/j.heliyon.2023.e17486)
9. Sorour AA et al. Vascular surgery integrated resident selection criteria in the pass or fail era. *J Vasc Surg*. 2023;77(2):625-631.e8. [doi:10.1016/j.jvs.2022.08.012](https://doi.org/10.1016/j.jvs.2022.08.012)
10. Ozair A, Bhat V, Detchou DKE. The US Residency Selection Process After the United States Medical Licensing Examination Step 1 Pass/Fail Change: Overview for Applicants and Educators. *JMIR Med Educ*. 2023;9:e37069. [doi:10.2196/37069](https://doi.org/10.2196/37069)
11. Zhang RV, Awan OA, Resnik CS, Hossain R. Potential Impact of a Pass or Fail United States Medical Licensing Exam Step 1 Scoring System on Radiology Residency Applications. *Acad Radiol*. 2022;29(1):158-165. [doi:10.1016/j.acra.2020.10.016](https://doi.org/10.1016/j.acra.2020.10.016)
12. Kim DY et al. General Surgery Residents can Be a Reliable Resource in the Evaluation of Residency Applications. *J Surg Educ*. 2015;72(6):e172-e176. [doi:10.1016/j.jsurg.2015.08.006](https://doi.org/10.1016/j.jsurg.2015.08.006)
13. Harrison LE. Using holistic review to form a diverse interview pool for selection to medical school. *Bayl Univ Med Cent Proc*. 2019;32(2):218-221. [doi:10.1080/08998280.2019.1576575](https://doi.org/10.1080/08998280.2019.1576575)
14. Conrad SS, Addams AN, Young GH. Holistic Review in Medical School Admissions and Selection: A Strategic, Mission-Driven Response to Shifting Societal Needs. *Acad Med*. 2016;91(11):1472-1474. [doi:10.1097/ACM.0000000000001403](https://doi.org/10.1097/ACM.0000000000001403)
15. Nakae S et al. Enrollment Management in Undergraduate Medical School Admissions: A Complementary Framework to Holistic Review for Increasing Diversity in Medicine. *Acad Med*. 2021;96(4):501-506. [doi:10.1097/ACM.0000000000003866](https://doi.org/10.1097/ACM.0000000000003866)
16. Wolfson RK et al. Residency Program Directors' Views on Research Conducted During Medical School: A National Survey. *Acad Med*. 2023;98(10):1185-1195. [doi:10.1097/ACM.0000000000005256](https://doi.org/10.1097/ACM.0000000000005256)
17. Patel H, Yakkanti R, Bellam K, Agyeman K, Aiyer A. Innovation in Resident Selection: Life Without Step 1. *J Med Educ Curric Dev*. 2022;9:238212052210849. [doi:10.1177/23821205221084936](https://doi.org/10.1177/23821205221084936)
18. Choi P, Langenau E, Roberts M, Blalock TW. Perspectives of Dermatology Program Directors on the Impact of Step 1 Pass/Fail. *Cureus*. Published online March 2023. [doi:10.7759/cureus.35801](https://doi.org/10.7759/cureus.35801)
19. Raborn LN, Janis JE. Current Views on the New United States Medical Licensing Examination Step 1 Pass/Fail Format: A Review of the Literature. *J Surg Res*. 2022;274:31-45. [doi:10.1016/j.jss.2022.01.002](https://doi.org/10.1016/j.jss.2022.01.002)
20. Althans AR et al. Exploring Characteristics of Academic General Surgery Residency Applicants: A Group Concept-Mapping Approach. *J Surg Educ*. 2022;79(6):1342-1352. [doi:10.1016/j.jsurg.2022.06.005](https://doi.org/10.1016/j.jsurg.2022.06.005)

21. Vaughan LA, Quick JA. Evidence-Based Selection of Surgical Residents. *Surg Clin North Am*. 2021;101(4):667-677. [doi:10.1016/j.suc.2021.05.012](https://doi.org/10.1016/j.suc.2021.05.012)