



Utilization of a Novel Program-Specific Evaluation Tool Results in a Decidedly Different Interview Pool Than Traditional Application Review

Joshua Lyons, MD, Katherine Bingmer, MD, John Ammori, MD, and Jeffrey Marks, MD

University Hospitals Cleveland Medical Center, Cleveland, Ohio

BACKGROUND: There are almost twice as many applicants as there are general surgery internships, each utilizing a common application with standard components. These elements are frequently not useful in determining affinity for a program or overall ability, and resultant poor fit may be partially responsible for program attrition. Alternative evaluation instruments would be beneficial to both programs and applicants.

METHODS: An application review committee comprised of resident representatives, faculty representing all program-affiliated institutions, and program leadership completed a written evaluation developed by a third party (SurgWise Consulting) that specializes in industrial and organizational psychology. The responses were compiled to create a standardized assessment tool. This assessment was sent to applicants who were subsequently ranked according to fit with our program. The pool of applicants was separately evaluated using our traditional application review. Two residents independently graded each applicant on a 5-point Likert scale to evaluate common application elements; applicants were subsequently assigned an overall score.

RESULTS: The assessment was completed by 507 (99%) of 512 qualifying applicants. Separately, 378 applications were reviewed by the traditional method for a total of 756 reviews. Of the 96 applicants identified by the assessment tool to invite for interviews, 22 (23%) qualified for interview invitations according to the traditional review method. The assessment produced 74 applicants that otherwise would not have been interviewed.

CONCLUSION: Traditional application review strategies have many shortcomings. A competency-based assessment

tool in the residency application selection process identifies a pool of applicants not identified by traditional review methods. (J Surg Ed 76:e110–e117. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: assessment, structured interviews, rank list, match

COMPETENCIES: Interpersonal and Communication Skills, Professionalism

INTRODUCTION

General surgery is a competitive field with a rigorous selection process. Programs attract many talented applicants with competitive resumes.¹ Every year there are significantly more applicants than available positions² and programs receive hundreds of applications for relatively few positions.¹ The applicant selection process utilized by many programs is costly, time consuming, and imperfect. With an overall attrition rate of around 13%,³ the stakes are high to select residents with a good fit for their prospective program.

When surveyed, program directors frequently report using the same few factors when selecting residents for interview.⁴ United States Medical Licensing Exam (USMLE) scores and letters of recommendation are most likely to have an impact on the global perception of an applicant's overall score.⁵ Despite the widespread use of application components to select applicants for interview, the objective measures most frequently used to differentiate potential residents have been shown to have little reliability to predict which residents will excel.^{6,7} USMLE scores have been shown to correlate with American Board of Surgery In-Training Examination scores, but are not predictive of overall resident performance.⁸

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Correspondence: Inquiries to Jeffrey M. Marks, MD, 11100 Euclid Avenue, Cleveland, OH 44106-5047; e-mail: jeffrey.marks@uhhospitals.org

When surveyed, 80% of orthopedic surgery program directors report using minimum USMLE Step 1 score, and frequently cited an interest in reducing the high volume of applications as the reason for doing so.⁹ The use of a USMLE Step 1 minimum score is a practice utilized at our program, and common throughout the country; 94% of surveyed program directors used USMLE Step 1 score in selecting applicants to interview, and it was the highest rated of all factors.⁴ A large number of applicants are being automatically rejected due to a test score that has not been shown to correlate with future performance as a resident.

Utilization of an online assessment tool to select applicants with a good fit for fellowship programs has been described.¹⁰ We instituted a similar process at our program, with the goal of simplifying applicant selection, selecting applicants with good fit for our program, and reducing the number of interviews required to have a successful match. Our hypothesis was that an online competency based assessment tool would identify different applicants to interview than traditional review methods.

METHODS

For the 2019 application and interview season, 2 methods of applicant review and assessment were used. The first was using a resident review process (traditional review) that has been utilized at our program for many years and the other using an online competency-based assessment tool sent to selected applicants.

For the traditional review process, initial filters were applied to the entire applicant group. These included a minimum USMLE Step 1 score of 230, anticipated graduation of a medical school located in the United States with an MD degree, and no previous felonies. Every applicant on the resulting list was then randomly assigned to 2 current residents, excluding current PGY-1 residents. The assigned residents then scored the applicants in 6 categories, each using a 5-point Likert scale.

Categories included overall rank, grades/scores, letters of recommendation, research/scholarly activity, personal statement, and extracurricular activities. The applicant's application, as accessible through the Program Director's Workstation website provided by the Electronic Residency Application Service, was the sole source of information for this review. If there was a greater than 2-point difference in the overall rank of an applicant between the 2 reviewers, the applicant was assigned to a third reviewer. To account for variability between reviewers, including giving persistently high or low scores, each score given for each category was then standardized for that reviewer with a resultant z-score. A z-score is a calculation that provides a standardized score that in our case corrects for individual reviewers that tend to score higher or lower than the average. A weighted average was used to produce a single numerical score per applicant per reviewer with the highest weight being attributed to overall score. Those resulting scores were then averaged between the 2 reviewers (or 3 if needed), producing a final numerical ranking. The top 120 applicants as produced by this score would have been invited to interview in previous years.

For the online assessment tool selection process, initial filters again were applied. These filters were identical to those used for the resident review process, but instead instituted a minimum USMLE Step 1 score of 210. Using a third-party consulting firm that specializes in surgery trainee selection (SurgWise Consulting), an assessment tool was developed. This competency-based tool is designed to emphasize fit and fairness to better assess an applicant's chances to succeed in a certain residency training program. After conducting a rigorous on-site job analysis, the consulting firm identified the most desired and required competencies needed upon entry for interns to succeed in the program and align with organizational values. A large number of situational judgment test items were developed to evaluate these competencies (Fig. 1). Review and input from a predefined group of residents and attendings provided input to the

You are a resident that frequently interacts with ER attendings. Unfortunately, the Emergency Medicine department and the prior resident rotating on the service had a history of confrontation. As a result, you find that your actions are constantly being scrutinized by ER faculty. Your attending has received multiple complaints about your interpersonal behavior, although you are certain they are unfounded. Please indicate the effectiveness of each of the following responses you could take next:

	Not effective at all	Slightly effective	Moderately effective	Very Effective	Extremely effective
Tell your attending the complaints are without merit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speak with the EM faculty to inquire more about how your behavior is being perceived	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apologize to EM faculty and monitor your behavior closely	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do nothing, and keep to yourself when in the ED	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speak with the Fellowship Director about these issues	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

FIGURE 1. An example situational judgment test, similar to those given by the online assessment.

assessment as part of the validation process. The resulting assessment tool was sent to the applicants meeting the above filters. Responses were analyzed by comparing the applicant's responses to those of our residents and attendings, and the top 91 applicants of the resulting ranked list were invited for an interview. In addition, 5 internal applicants were also interviewed who did not perform high enough on the assessment to otherwise qualify for an interview. This brought the total number of interviews invitations to 96. Several interview offers were declined, and offers were subsequently extended on a rolling basis to applicants who were on the waiting list. The reduction in applicants invited to interview was based on the applicants having already been evaluated by the assessment tool and therefore have a higher likelihood of interviewing well with our program which would allow for a shorter rank list.

Previously, interviews were unstructured with no guidance given to the interviewers and each interviewer giving overall numerical score for each applicant. In addition to the assessment tool, the same consulting team held a 4-hour workshop for our interviewers, training them on how to conduct an effective and fair interview. Each interviewer asked applicants 3 scenario-based questions which covered topics such as dependability, teamwork, perseverance, adaptability, integrity, and problem solving. The structured interviews were tailored to the competencies previously determined to be valued by the program during the on-site visit. The structured interview was comprised of behavioral-focused, experience-based questions, as this style of questioning has shown to have job performance validity.¹¹ In addition, each interviewer scored each applicant on their communication skills, motivational fit, and interpersonal skills. Each interview resulted in 6 scores and each applicant had 3 separate interviews for a total of 18 scores per applicant. This interview process provides a more comprehensive and impartial assessment of the applicants and has been shown to better predict future job performance in corporate settings.¹² To again account for variability between interviewers including interviewers giving persistently high or low scores, each score was standardized for that reviewer producing a *z*-score. The resulting *z*-scores were averaged to produce a final numerical ranking which produced a preliminary rank list. Because all interviewed applicants showed high performance on the online competency-based assessment tool, only the interview scores were used to form the preliminary rank list.

A rank list meeting was then held; all categorical residents, interviewers, and other members of the Department of Surgery were invited. Slight changes were made to the rank list if the majority of the meeting's participants reached a consensus and the top 71 applicants were ranked.

Statistical analyses were done with JMP Pro 13 (SAS Institute Inc, Cary, NC) and used the Spearman's rank correlation coefficient to compare the similarity between ranked lists with a *p* value of <0.05 being considered significant.

RESULTS

Our program received 1163 total applications. After applying the filters used for the traditional review process, 378 applicants were identified and reviewed. Applying the filters used for the online assessment tool, 512 applicants were identified and sent the assessment. Therefore, an additional 134 applicants (35% more than the traditional review) were evaluated using the assessment tool. Five hundred and seven of the 512 applicants completed the assessment for a 99% response rate.

Of the 96 applicants identified by the assessment tool to invite for interviews, 22 (23%) of them also qualified for interview invitations according to the traditional review method. Therefore, the assessment produced 74 applicants that otherwise would not have been interviewed (Fig. 2). In addition, 21 (28%) of the 74 applicants that otherwise would not have been interviewed had USMLE Step 1 scores below 230 and therefore were not reviewed by the traditional method.

There were no significant differences in the average age, gender, number of underrepresented minorities, or medical school type (public vs private) between applicants identified for an interview by the traditional review and those identified by the online assessment. Due to the lower USMLE Step 1 score filter used by the online assessment, there was a statistically significant difference in average USMLE Step 1 score. ($p < 0.05$ was used to determine statistical significance) (Table 1).

Because the only difference in filters between the 2 review methods was USMLE Step 1 score, every applicant reviewed by the traditional review method was also sent an online assessment and therefore the 2 ranked lists can be compared for similarity, with the exception of those applicants who did not meet the USMLE step 1 score filter for traditional review. The 2 resulting ranked lists were compared and were not statistically similar with an r_s value of 0.03 ($p = 0.51$). (Fig. 3).

Of the 71 applicants ranked, 21 (30%) of them would possibly also have been ranked if the traditional review method was used as they also would have been offered an interview. Therefore, the assessment produced an additional 50 applicants that were eventually ranked that initially would not have been offered an interview according to the traditional review method. (Fig. 4). In addition, 17 (24%) of the applicants ranked had USMLE step 1 scores below 230 and therefore were not reviewed by the traditional method.

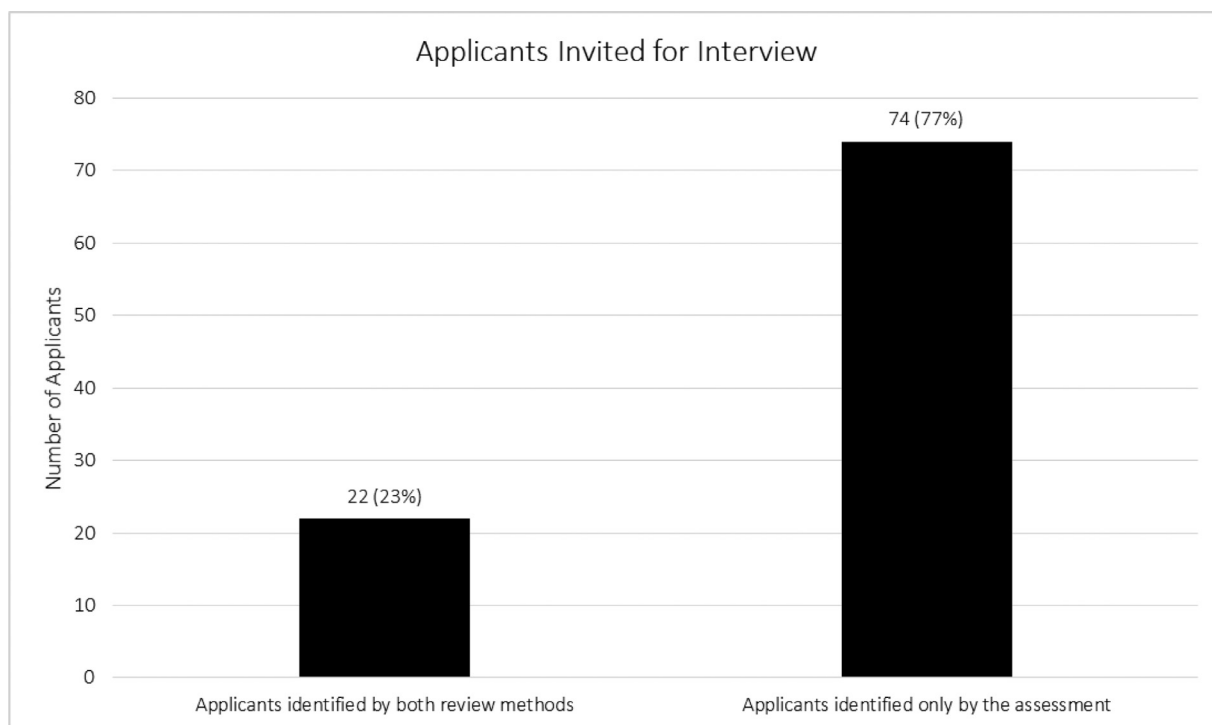


FIGURE 2. How the 96 applicants interviewed were identified for an interview.

TABLE 1. Demographic Comparison of Applicants Identified for an Interview by the Traditional Review Method to those Identified by the Online Assessment

	Traditional Review	Online Assessment	p Value
Age (y)	27.2	27.5	0.32
Female gender	53 (55.2%)	55 (57.3%)	0.77
USMLE score	246	239	<0.01
URM	8	7	0.79
US public school	58 (60.4%)	55 (57.3%)	0.66
US private school	38 (39.6%)	41 (42.7%)	

URM, "Underrepresented minorities" according to the AAMC definition.²⁸

In order to compare the results of the online assessment tool and the results of the interview process, the ranked lists from both the online assessment tool and the interview process were compared for similarity and were found to be not statistically similar with an r_s value of -0.02 ($p = 0.83$).

Of the 7 matched PGY-1 residents, none would have been offered an interview if the traditional review method was used. This is partially due to the leniency in USMLE Step 1 minimum score, as 2 of the 7 applicants would not have met the minimum score needed for inclusion in the traditional interview process.

According to the National Resident Matching Program (NRMP) match result data, general surgery residency programs in 2019 needed to rank 5.0 applicants per position to fill all positions.² Our program currently has 7 available PGY-1 positions and therefore on average, we would match down to position 35 on our rank list. For the 2019 match, we matched down to position 19 on our rank list, a significant improvement to the average of 35. Using our match data for the previous 5 years, on average we matched down to position 50 on our rank list.

DISCUSSION

The online assessment tool was completed by 99% of applicants, which is significantly higher than other survey types and can be used to specifically evaluate a large group of applicants.¹³ The assessment tool resulted in a decidedly different interview selection pool compared to traditional methods. While there were no significant differences in age, gender, or underrepresented minorities, 77% of applicants would not have been considered by traditional application review. The online assessment tool did use a lower USMLE Step 1 score in the initial filtering, but this only accounted for 28% of those applicants that otherwise would not have been interviewed if the traditional method was used. Therefore, the majority of invited applicants that otherwise would not have

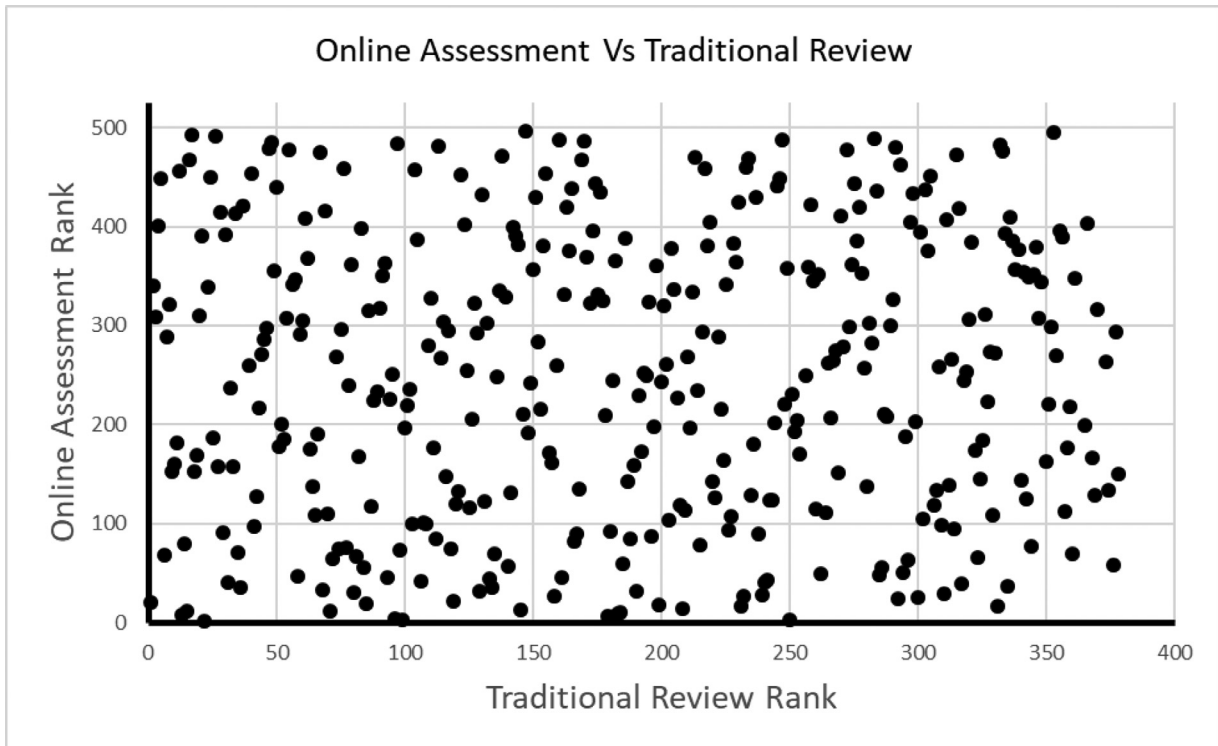


FIGURE 3. A comparison between the online assessment tool rank and traditional review rank showing no correlation. Each dot represents one applicant.

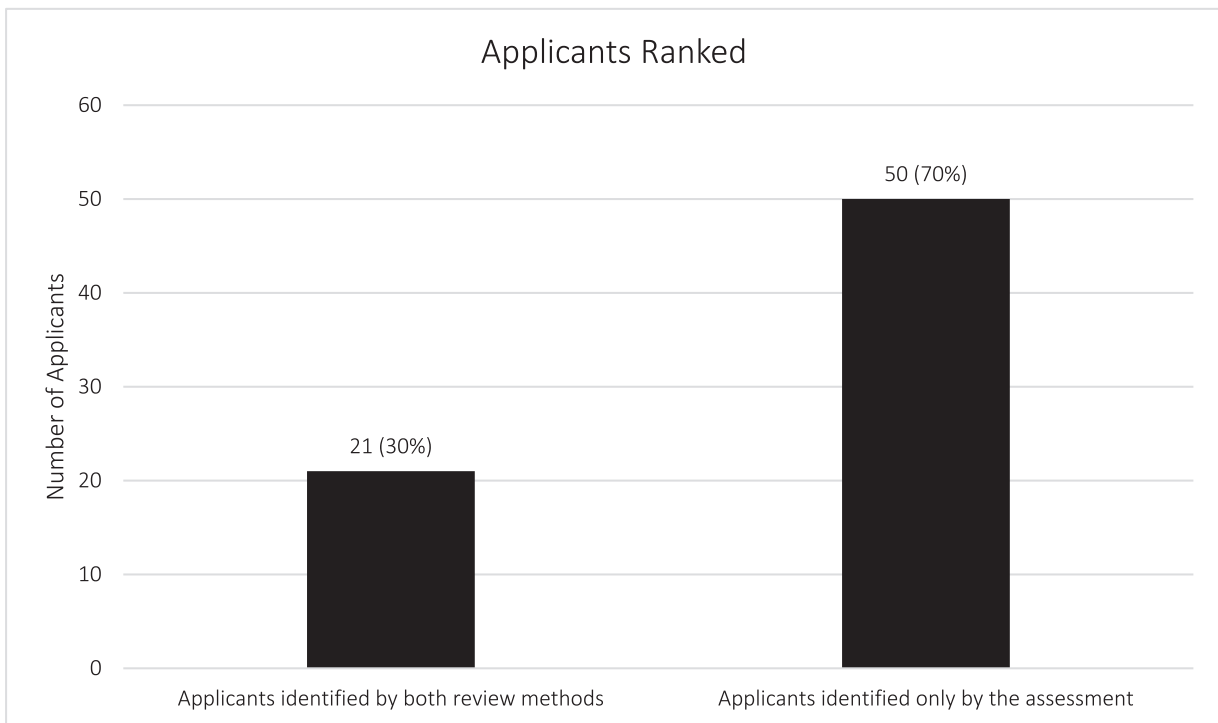


FIGURE 4. How the 71 applicants ranked were initially identified for an interview.

been invited if the traditional method was used were selected based on their performance on the online assessment and not due to the lower USMLE Step 1 filter. The final rank list was also distinctly different from a rank list that would have been produced if traditional methods were used. Our last applicant matched was much higher on our rank list than in previous years. Because the rank list was a reflection of structured interview performance, matching applicants higher on our rank list indicates our matched applicants received high scores in their structured interviews. Because structured interviews have been shown to predict performance on ACMGE milestones, overall resident performance, and potentially decrease surgical resident attrition, our matched applicants should perform well in our program and will be studied in the future.¹⁴⁻¹⁶

Not only did the online assessment tool identify a different pool of applicants to invite for an interview, there was no correlation with how applicants performed on the online assessment tool and how they were scored based on the traditional review method. In addition, despite the online assessment tool being formed by evaluating our own residents and attendings, there was no correlation between performing well on the online assessment and performing well in the structured interview. The disparity in scoring between the preinterview online assessment and the structured interview scoring suggests that they are testing different components of an applicant's skills sets.

Finally, all matched PGY-1 residents were identified and interviewed based on the online assessment tool; none would have been interviewed if the traditional method was used. In addition, we matched significantly higher on our match list than the national average and much higher than our previous match results. This potentially is a result of the fact that the competency-based assessment tool and structured interviews lead to finding applicants who are a better fit for our program, although true fit is not able to be assessed until resident performance can be evaluated. This may then result in needing to interview fewer applicants, as we did this year, if this model is followed which can then lead to substantial cost savings for residency programs.¹⁷

The task of sorting through hundreds of applicants can be a daunting one, and frequently program directors use USMLE scores to eliminate a large pool of applicants.⁹ This results in a large portion of the applicant pool being excluded. Test scores have not been shown to correlate to resident performance,⁷ the grading of personal statements has been shown to lack objectivity,¹⁸ and the more subjective components of the application are not reliable assessments. While letters of recommendation are frequently weighted heavily, they can be worded differently based on gender.^{5,19} Program directors consider the importance

of research experience to be secondary to test scores and letters of recommendation.²⁰ While imperfect, the traditional method of evaluating applications is a reflection of program directors doing the best they can with the information they are given.

Attempts to improve applicant selection have largely focused on the interview process. A working interview, where applicant spend the day with a surgical team and are observed in clinical situations, has been described.²¹ Other methods of nonconventional interviews that have been trialed include group interviews and skills assessments to better assess resident potential and fit.^{22,23} One study even developed a surgery-specific exam that was administered to interviewing residents and found to correlate better to American Board of Surgery In-Training Examination scores than USMLE Step 1 or 2.²⁴ Structured interviews are largely regarded as the superior interview method and have previously shown to reduce attrition in surgical residency, although few programs utilize these methods.^{16,25,26} However, all of these interventions occur at the time of the interview. There have been no attempts in the literature to improve upon the selection of resident applicants invited to interview.

The applicant selection process is costly for general surgery residency programs with an average cost of over \$100,000 as many hours are dedicated by administrators, faculty, and residents.¹⁷ Much of the cost is incurred from the interviews, and a large amount of the cost is directly related to the number of applicants interviewed by a program.¹⁷ Faculty interviews are often a good assessment of applicant fit and future performance.⁸ Inter-rater reliability is frequently a concern, but can be improved with structured interviewing; a 1 day course has been shown to improve inter-rater consistency in the evaluation of a surgical applicant.²⁷ The interview process is very important in the selection of applicants, but fewer interviews cut down on cost and administrative burden. In our study, we decreased our number of interviews from 120 to 96, and still matched 7 out of our top 19 applicants. According to national survey data obtained from general surgery program directors, the average cost to interview an applicant is \$1221 and therefore our reduction in the number of interviews results in an almost \$30,000 cost savings for the department.¹⁴ The reduced spending in interview season must be balanced with the additional cost pertaining to the consulting firm, however, utilization of the assessment tool over multiple interview seasons would improve cost-effectiveness. In the future, this evaluation tool could allow us to reduce the numbers of interviews further and therefore further reduce the overall cost of the applicant selection process.

There are several limitations to our study. This is a single-center experience and this type of applicant review to our knowledge as not been used to evaluate a large number of

residency applicants. In addition, while interviewers did undergo a structured interview training workshop, they were certainly inexperienced with this new type of interview process and therefore the structured interviews were likely not utilized to their full potential. As both the online assessment and the structured interviews were implemented in the same application year, we cannot know the effect of each on the resulting improved match list, although certainly both played a role. Finally, true fit of applicants within the program cannot be assessed until resident performance and attrition data can be collected.

Future areas of study include comparing residents that matched using the traditional review process to those who just matched using the online assessment and structured interview process. Residents will be compared based on their performance on both standardized exams and internal review in addition to their subjective residency experience.

This study demonstrates that an online evaluation tool is a feasible way to gather additional information about general surgery applicants. Using this tool, we were able to consider a larger number of applicants for interview selection. To the best of our knowledge, this is the first study to attempt to improve the resident applicant selection process at the preinterview stage. The online competency assessment tool resulted in a decidedly different interview pool, rank list, and matched interns than our traditional method.

REFERENCES

1. Dort JM, Trickey AW, Kallies KJ, Joshi ART, Sidwell RA, Jarman BT. Applicant Characteristics associated with selection for ranking at independent surgery residency programs. *J Surg Educ*. 2015. <https://doi.org/10.1016/j.jsurg.2015.04.021>.
2. Program NRM. *National Resident Matching Program, Results and Data: 2018*. Main Residency Match. Washington, DC.
3. Shweikeh F, Schwed AC, Hsu CH, Nfonsam VN. Status of resident attrition from surgical residency in the past, present, and future outlook. *J Surg Educ*. 2018. <https://doi.org/10.1016/j.jsurg.2017.07.015>.
4. Program NRM. *National Resident Matching Program, Data release and research committee: results of the 2018 NRMP Program Director Survey*. Washington, DC; 2018.
5. Sharp C, Plank A, Dove J, et al. The predictive value of application variables on the global rating of applicants to a general surgery residency program. *J Surg Educ*. 2015. <https://doi.org/10.1016/j.jsurg.2014.06.003>.
6. Hughes BD, Perone JA, Cummins CB, et al. Personality testing may identify applicants who will become successful in general surgery residency. *J Surg Res*. 2019. <https://doi.org/10.1016/j.jss.2018.08.003>.
7. Fryer JP, Corcoran N, George B, Wang E, Darosa D. Does resident ranking during recruitment accurately predict subsequent performance as a surgical resident? *J Surg Educ*. 2012. <https://doi.org/10.1016/j.jsurg.2012.06.010>.
8. Brothers TE, Wetherholt S. Importance of the faculty interview during the resident application process. *J Surg Educ*. 2007. <https://doi.org/10.1016/j.jsurg.2007.05.003>.
9. Schrock JB, Kraeutler MJ, Dayton MR, Mccarty EC. A cross-sectional analysis of minimum usmle step 1 and 2 criteria used by orthopaedic surgery residency programs in screening residency applications. *J Am Acad Orthop Surg*. 2017. <https://doi.org/10.5435/JAAOS-D-16-00725>.
10. Gardner AK, Dunkin BJ. Pursuing excellence: the power of selection science to provide meaningful data and enhance efficiency in selecting surgical trainees. *Ann Surg*. 2018.
11. Pulakos ED, Schmitt N. Experience-based and situational interview questions: studies of validity. *Pers Psychol*. 1995. <https://doi.org/10.1111/j.1744-6570.1995.tb01758.x>.
12. CM A, Pursell ED, Brown BK. Structured interviewing: raising the psychometric properties of the employment interview. *Pers Psychol*. 1988. <https://doi.org/10.1111/j.1744-6570.1988.tb00630.x>.
13. Baruch Y, Holtom BC. Survey response rate levels and trends in organizational research. *Hum Relations*. 2008. <https://doi.org/10.1177/0018726708094863>.
14. Marcus-Blank B, Dahlke JA, Braman JP, et al. Predicting performance of first-year residents. *Acad Med*. 2019. <https://doi.org/10.1097/acm.0000000000002429>.
15. Gardner AK, Dunkin BJ. Evaluation of validity evidence for personality, emotional intelligence, and situational judgment tests to identify successful residents. *JAMA Surg*. 2018. <https://doi.org/10.1001/jamasurg.2017.5013>.
16. Kelz RR, Mullen JL, Kaiser LR, et al. Prevention of surgical resident attrition by a novel selection strategy. *Ann Surg*. 2010. <https://doi.org/10.1097/SLA.0b013e3181f27a50>.
17. Gardner AK, Smink DS, Scott BG, Korndorffer JR, Harrington D, Ritter EM. How much are we

- spending on resident selection. *J Surg Educ*. 2018. <https://doi.org/10.1016/j.jsurg.2018.10.001>.
18. 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. <https://doi.org/10.1016/j.jsurg.2011.12.003>.
 19. Turrentine F, Dreisbach C, St Ivany A, Hanks J, Schroen A. Influence of gender on surgical residency applicants' recommendation letters. *J Am Coll Surg*. 2019;12:356-365.
 20. Melendez MM, Xu X, Sexton TR, Shapiro MJ, Mohan EP. The importance of basic science and clinical research as a selection criterion for general surgery residency programs. *J Surg Educ*. 2008. <https://doi.org/10.1016/j.jsurg.2007.08.009>.
 21. Seabott H, Smith RK, Alseidi A, Thirlby RC. The surgical residency interview: a candidate-centered, working approach. *J Surg Educ*. 2012. <https://doi.org/10.1016/j.jsurg.2012.06.004>.
 22. Ogunyemi D, Alexander C, Tangchitnob E, Kim DS. Mini surgical simulation, role play, and group and behavioral interviews in resident selection. *J Grad Med Educ*. 2016. <https://doi.org/10.4300/jgme-d-15-00203.1>.
 23. Panait L, Larios JM, Brenes RA, et al. Surgical skills assessment of applicants to general surgery residency. *J Surg Res*. 2011. <https://doi.org/10.1016/j.jss.2011.04.006>.
 24. 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. <https://doi.org/10.1016/j.jsurg.2012.05.011>.
 25. Kim RH, Gilbert T, Suh S, Miller JK, Eggerstedt JM. General surgery residency interviews: are we following best practices. *Am J Surg*. 2016. <https://doi.org/10.1016/j.amjsurg.2015.10.003>.
 26. Joshi ART, Vargo D, Mathis A, Love JN, Dhir T, Termuhlen PM. Surgical residency recruitment—opportunities for improvement. *J Surg Educ*. 2016. <https://doi.org/10.1016/j.jsurg.2016.09.005>.
 27. Gardner AK, D'Onofrio BC, Dunkin BJ. Can we get faculty interviewers on the same page? An examination of a structured interview course for surgeons. *J Surg Educ*. 2018. <https://doi.org/10.1016/j.jsurg.2017.06.006>.
 28. AAMC Executive Committee. Underrepresented in medicine definition - initiatives - AAMC. Retrieved May 31, 2019. <https://www.aamc.org/initiatives/urm/>.

SUPPLEMENTARY INFORMATION

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.jsurg.2019.10.007](https://doi.org/10.1016/j.jsurg.2019.10.007).