

Innovation in Resident Selection: Life Without Step 1

Hares Patel¹, Ram Yakkanti², Krishna Bellam³, Kofi Agyeman²
and Amiethab Aiyer²

¹University of Kentucky, College of Medicine, USA. ²Department of Orthopedic Surgery, University of Miami, USA. ³Warren Alpert Medical School of Brown University, USA.

Journal of Medical Education and
Curricular Development
Volume 9: 1–11
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23821205221084936



ABSTRACT

INTRODUCTION: The announcement of Step 1 shifting to a Pass/Fail metric has prompted resident selection committees (RSCs) to pursue objective methods of evaluating prospective residents. Regardless of the program's specialty or affiliated hospital/school, RSCs universally aim to recognize and choose applicants who are an "optimal fit" to their programs.¹ An optimal fit can be defined as a candidate who thrives in the clinical and academic setting, both contributing to and benefiting from their respective training environments.

OBJECTIVE: The objective of this scoping review is to evaluate alternative, innovative methods by which RSCs can evaluate applicants and predict success during residency. Objective methods include: Step 2 scores, Traditionally Used Metrics (core clerkship scores), interview performance, musical talent, sports involvement, AOA membership, research publications, unprofessional behavior, Dean's letters, Rank list, judgement testing, and specialty-specific shelf exams.^{13–15}

METHODS: A scoping review was performed in compliance with the guidelines indicated by the PRISMA Protocol for scoping review.¹⁸ 9308 results were identified in the original PubMed search for articles with the key words "Resident Success". Abstract screening and application of inclusion and exclusion criteria yielded 97 articles that were critically appraised via review of full manuscript.

RESULTS: Of the articles that focused on personality traits, situational judgement testing, and specialty specific pre-assessment, all of them demonstrated some level of predictability for resident success. Standardized Letter of Recommendations, Traditionally Used Metrics, and STEP 2 did not show a unanimous consensus in demonstrating predictability of a resident's success, this is because some articles suggested predictability and some articles disputed predictability.

CONCLUSION: The authors found personality traits, situational judgement testing, and specialty specific assessments to be predictive in selecting successful residents. Further research should aim to analyze exactly how RSCs utilize these assessment tools to aid in screening their large and competitive applicant pools to find residents that will be successful in their program.

KEYWORDS: residency, success, application, USMLE, step 1

RECEIVED: November 11, 2021. **ACCEPTED:** January 25, 2021

TYPE: Review

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Hares Anilkumar Patel, University of Kentucky, College of Medicine, USA.
Email: Hares.patel@uky.edu

Introduction

The announcement of Step 1 shifting to a Pass/Fail metric has prompted resident selection committees (RSCs) to pursue other objective methods of evaluating prospective residents. Regardless of the program's specialty or affiliated hospital/school, RSCs universally aim to recognize and choose applicants who are an "optimal fit" to their programs.¹ An optimal fit can be defined as a candidate who thrives in the clinical and academic setting, both contributing to and benefiting from their respective training environments. More specifically, the objectives of RSCs are to match medical students who will execute training tasks well, not only in the residency-specific hospital setting, but also in their ultimate career paths.^{1,2} This is a challenging endeavor likely to become more so with the advent of the Pass/Fail Step 1 System.

With the numerical Step 1 score being nulled, there are several downsides that have been discussed in literature. The United States Medical Licensing Examination (USMLE) offers a highly reliable, objective assessment of relevant

competencies in a nationally standardized context. Furthermore, due to a majority of Medical Student Performance Evaluation (MSPEs) and medical school courses adopting a pass/fail system, there is a notable paucity of useful objective assessment tools; some more arbitrary ranking systems at times signify "outstanding" performances as meaning first or fourth quartile.³ Therefore, without such useful information in MSPEs and graded course performance, numeric Step 1 scores emerged as critical evaluation of applicants for residency. The Association of Program Directors in Radiology (APDR) is concerned that pass/fail reporting of the USMLE Step 1 score would take away an objective measure of medical student's knowledge and the incentive to acquire as much of it as possible.⁴ Deng et al further supports this idea, stating that Step 1 scores were strongly and positively associated with performance on in-training and specialty board certification exams, along with subjective faculty evaluations of resident clinical performance.⁵ More specifically, there is a correlation between American Board of Orthopedic Surgery (ABOS) pass



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

rate and Step 1/Step 2 scores, as the pass rate for those who score above a 220 and 240 is 96% and 100%, respectively.⁶⁻⁸

That being said, while there is evidence that USMLE correlate strongly with performance on similar licensing exams and other multiple-choice test scores; these measurements do not extrapolate as they lack association with measures of clinical skills that matter among advanced medical students, residents, and subspecialty fellows.⁹ Said a different way, there is no proven or published correlation between one's step 1 score and future patient outcomes, or ability to become a successful clinician. Furthermore, shifts to pass/fail scoring reduce student stress and reduce burnout, with concomitant improvement in student physiologic, mental, and behavioral health.^{2,10,11} With this, the time and effort devoted to preparing for Step 1 can be diverted to alternative activities that will prepare them to be good physicians, resulting in making them more prepared to provide quality care, improve their residency performance, and ultimately improve patient outcomes.^{1,2} Lastly, studies do note some differences in Step 1 scores attributable to race and ethnicity, with self-identified black, Asian, and Hispanic examinees showing score differences when compared with self-identified white examinees.¹²

The transition to a pass/fail system presents RSCs with the perfect opportunity to explore novel ways to determine best fits and predict performance. The objective of this scoping review is to evaluate alternative, innovative methods by which RSCs can evaluate applicants and predict success during residency. Objective methods include: Step 2 scores, Traditionally Used Metrics (core clerkship scores), interview performance, musical talent, sports involvement, Alpha Omega Alpha Honor Society (AOA) membership, research publications, unprofessional behavior, Dean's letters, Rank list, judgement testing, and specialty-specific shelf exams.¹³⁻¹⁵ Related subjective assessments include standardized letters of recommendations (LORs) and personality traits.^{16,17}

Methods and Materials

A systematic review of the literature was performed in compliance with the guidelines indicated by the PRISMA Protocol for scoping review.¹⁸ A literature search was conducted on PubMed on Oct. 14th, 2020: the results were limited to papers in English and that were published on or after 2005. 2005 was selected as the authors determined that a 15-year span would be an appropriate range to encompass the most up to date practices in evaluating residency applicants. The search was conducted with the key word "Resident Success" to cast a broad net to include as many publications that aim to explore which application factors predict residency success as possible. 2 reviewers evaluated the titles and abstracts of all the publications identified. To maximize the consistency among the reviewers the same format was utilized to evaluate article abstracts. This initial screen narrowed the pool of accepted publications. In this initial exclusion phase the articles

were organized to fit into one of five categories of resident evaluation other than USMLE Step 1: Other Standardized Exam (STEP 2), Traditionally Used Metrics, Situational Judgement Testing, *Standardized Letters of Recommendation (LORs)*, Specialty Specific Pre-assessment, Personality Traits. 2 reviewers summarized the full texts of the remaining 97 articles for analysis and exclusion. The inclusion criteria are (1) minimum sample size of 20, (2) statistical significance of presented results, and (3) compares a medical student application attribute to performance in residency. Exclusion criteria: (1) Non English-publications, (2) survey based studies and (3) non Resident/US medical school cohorts. Once the publications meeting inclusion and exclusion criteria were identified, the authors used a chart template to facilitate data extraction. Data extracted from the full articles included the journal that published the article, type of study, level of evidence, the variables included, outcomes reported, conclusions of the article, statistical significance, and the overall predictability of resident success. Results were continuously negotiated and updated in an iterative process by the authors. Figure 1 shows the categorical approach of the search process and exclusion steps. The remaining articles after screening, and application of exclusion criteria were read in their entirety and critically appraised. On Nov. 23rd, 2020 the search strategy was performed again to review any new abstracts and full text articles added since initial review, and above mentioned screening process repeated.

Results

9308 results were identified in the original PubMed search for articles with the key words "Resident Success". Abstract screening and application of inclusion and exclusion criteria yielded 97 articles that were critically appraised via review of full manuscript. After the full-text review was performed and the exclusion criteria described in the methods and materials were applied, the reviewers chose to include 9 articles focused on STEP 2 as a predictor of resident performance. 18 articles were identified that evaluated traditionally used metrics such as clerkship honors/grades, AOA membership, research experience, dean's letter, letter of recommendations, rank of the school, sports involvement, and musical talent. 5 articles explored situational judgement testing as a possible predictor of resident performance. 2 articles discussed a standardized letter of recommendation as implemented in the Emergency Medicine Specialty. 1 article assessed the use of specialty specific pre-assessment exams for applicants. 6 articles looked at the use of personality traits to determine if a resident would be successful at their program. Of these final articles, 10 of them were multi-institutional and 3 of them were multi-specialty as well (in the STEP 2 group, SJT group, and Traditional group). All other articles were single institution, single specialty. The average sample size and the specialties included in each category can be summarized in Table 1. All

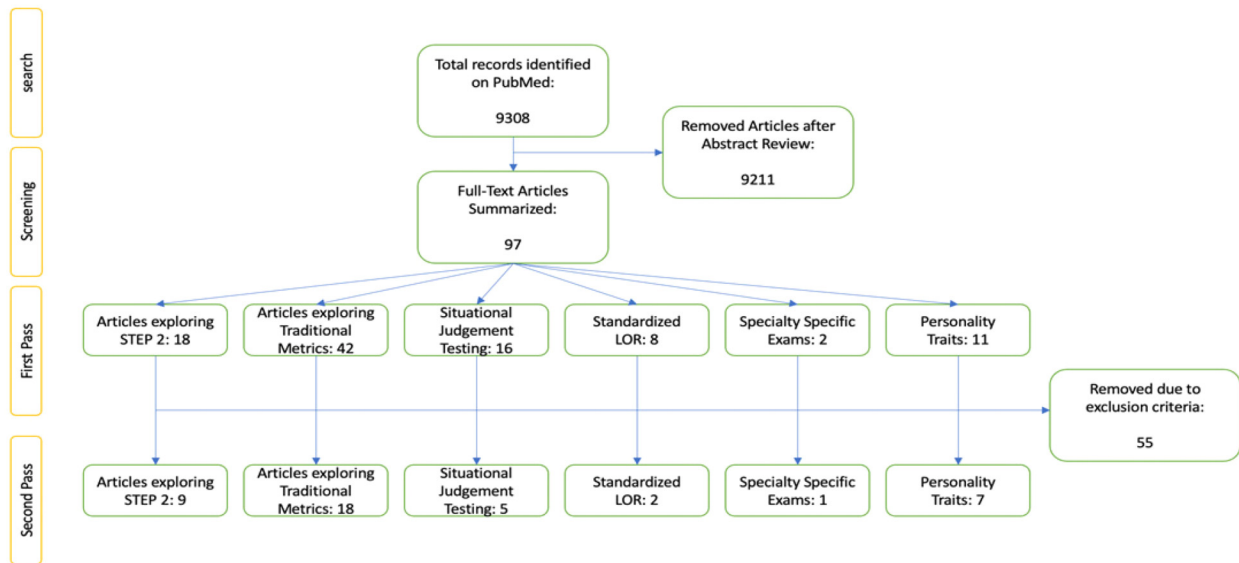


Figure 1. Scoping search and study flow diagram.

of the articles that passed the exclusion criteria in this study can be found in Table 2.

After the full text articles were summarized, they were labeled as predictive of resident success or not predictive of resident success. If the results of the article presented statistical significance and proposed the ability to predict a measurement of resident success, the article was deemed as “showing predictability”. Of the articles that focused on personality traits, situational judgement testing, and specialty specific pre-assessment, all of them demonstrated some level of predictability for resident success. Standardized Letter of Recommendations, Traditionally Used Metrics, and STEP 2 did not show a unanimous consensus in demonstrating predictability of a resident’s success, this is because some articles suggested predictability and some articles disputed predictability. Figure 2 depicts the comparison of articles that were evaluated within their respective categories. The figure demonstrates the number of articles in each category that showed predictability and the number of articles that were inconclusive in predicting resident success.

Discussion

STEP 1 shifting to a Pass/Fail (P/F) metric has prompted resident selection committees (RSCs) to investigate other objective methods of evaluating prospective residents. The shift in STEP 1 to P/F was largely justified as it addresses concerns of medical student well-being, racial inequity in STEP scores, and to allow students to focus more on developing clinical skills during their pre-clinical years. National surveys indicate that program directors across all specialties use STEP 1 scores to stratify their applicant pool. With the loss of STEP 1 scores, resident selection committees will need to find alternative methods to narrow the large applicant pool into a select number of applicants for an interview. In addition, there is concern that this change will put

some categories of students at a significant disadvantage. Introverts will be at a disadvantage as this change might make networking with faculty in the specialties they desire more important. In addition, international/DO students or students from “less prestigious” schools are at an inherent disadvantage because program directors might use school reputation as a method of screening the applicants.

This paper conducted a scoping review to investigate alternative methods through which resident selection committees may choose to evaluate and stratify applicants. The major themes found in this scoping review included USMLE STEP 2 scores, Traditionally Used Metrics, Situational Judgement Testing, Standardized LORs, Specialty Specific Pre-Assessment, and Personality Traits. The articles included in this paper covered specialties including anesthesia, psychiatry, orthopedics, general surgery, emergency medicine, ear, nose, throat surgery (ENT), dermatology, neurosurgery, internal medicine, pathology, PMnR, radiology, urology, and obstetrics and gynecology (OBGYN). Table 1 presents the breakdown for which specialties were included in specific article categories.

Step 2 and Traditionally Used Metrics

STEP 2 and other Traditionally Used Metrics are some of the first options residency programs are expected to use for stratifying applicants after the switch to P/F. STEP 2 was a major topic of analysis in 9 articles. 7 articles showed positive prediction of resident success, however, 2 of them disputed STEP 2 as a positive indicator of resident success. Traditionally Used Metrics were defined as clerkship honors/grades, AOA membership, research experience, dean’s letter, letter of recommendations, rank of the school, sports involvement, and musical

Table 1. Categories summary.

Category	
<i>STEP 2:</i>	
Total Articles Included	9
Predictive Article	6
Avg Sample Size	152
Specialties included	Orthopedics(3), Psychiatry(2), Anesthesiology(2), Gen Surg(4), EM, ENT, Derm, IM, Neuro, OBGYN, Pathology, PMnR
Multi-Institutional	1
<i>Traditionally Used Metrics:</i>	
Total Articles Included	19
Predictive Article	12
Avg Sample Size	167.5
Specialties included	Neurosurgery, ENT (2), Neurology, Radiology, Urology (4), orthopedic (2), EM (3), Gen Surg
Multi-Institutional Study Included	2
<i>Situational Judgment Testing:</i>	
Total Articles Included	5
Predictive Article	5
Avg Sample Size	85.4
Specialties included	Gen Surg, OBGYN(2), EM
Multi-Institutional Study Included	1
<i>Standardized LORs:</i>	
Total Articles Included	2
Predictive Article	1
Avg Sample Size	532
Specialties included	EM, ENT
Multi-Institutional Study Included	1
<i>Specialty Specific Pre-Assessment:</i>	
Total Articles Included	1
Predictive Article	1
Avg Sample Size	20
Specialties included	Orthopedics
Multi-Institutional Study Included	0

(continued)

Table 1. Continued.

Category	
<i>Personality Traits:</i>	
Total Articles Included	7
Predictive Article	7
Avg Sample Size	96.5
Specialties included	Orthopedics(3), Anesthesia, Gen Surg (5), Neurosurgery, Radiology, Urology(4), EM (4), ENT
Multi-Institutional Study Included	0

talent. 12 articles showed positive prediction, however, 7 articles showed negative prediction for resident success. Concerns of residency programs relying on STEP 2 scores and traditionally used metrics include continued pressure placed on medical students about performing well on one single test, or set of tests, that could have a huge impact on their ability to choose a residency. If program directors use STEP 2 scores and traditionally used metrics there is concern that the expected progression to more sophisticated methods of evaluating students will not occur, as was part of the intention for eliminating STEP 1 scores. The concern is that students will continue to struggle with intensified stress for higher scores and racial inequities associated with standardized tests will continue to persist. In addition, due to the inconsistency of STEP 2 and traditionally used metrics in predicting resident success it is hard to recommend STEP 2 and Traditionally Used Metrics as an innovative path in resident selection. Lastly, it is important to keep in mind that the articles included in this scoping review varied on their definition of resident success. Generally, the categories of “STEP 2” and “Traditionally Used Metrics” relied heavily on the use of specialty board exam pass rates as an indicator of resident success. Although it is highly important for residency programs to be confident that their residents will pass boards – this alone is not always a determinant of a successful resident or a future clinician.

Standardized Letter of Recommendation

Recently, Emergency Medicine introduced its own application to standardize how applicants apply for residency in EM. They include a Standardized Letter of Recommendation as part of the application (SLORs). In the review of articles that claim to determine if SLORs are effective in predicting resident success there was 1 that determined that SLORs were a predictive measure of resident success and 1 article that was not a predictive measure of resident success. Bhat et al found that “global rating” and “competitiveness” on non-program leadership

Table 2. Studies considered in this scoping review.

Category	1st Author	Article	PUBMED ID	Publication Year	Specialty Represented	Institutions Involved	Predictive
<i>STEP 2:</i>							
	Miller BJ	US Medical Licensing Exam scores and performance on the Psychiatry Resident In-Training Examination	24804630	2014	Psychiatry	Single-Institution	Yes
	Guffey RC	The utility of pre-residency standardized tests for anesthesiology resident selection: the place of United States Medical Licensing Examination scores	21048098	2011	Anesthesiology	Single-Institution	Yes
	Mainthia R	Achievement in surgical residency: are objective measures of performance associated with awards received in final years of training?	24602705	2014	Gen Surg	Single-Institution	No
	Harmouche E	USMLE Scores Predict Success in ABEM Initial Certification: A Multicenter Study	28435509	2017	EM	Multi-Institutional	Yes
	Puscas L	USMLE and Otolaryngology: Predicting Board Performance	28418270	2017	ENT	Single-Institution	No
	Cohen ER	Are USMLE Scores Valid Measures for Chief Resident Selection?	32879684	2020	Multi-Specialty	Single-Institution	No
	Al Fayyadh MJ	Predicting Success of Preliminary Surgical Residents: A Multi-Institutional Study	27395396	2016	Gen Surg	Single-Institution	Yes
	Dyrstad BW	Predictive measures of a resident's performance on written Orthopedic Board scores	22096449	2011	Orthopedic	Single-Institution	Yes
	Maker VK	Can we predict which residents are going to pass/fail the oral boards?	23111034	2012	Gen Surg	Single-Institution	Yes
<i>Traditionally Used Metrics:</i>							
	Zuckerman SL	Predicting Resident Performance from Preresidency Factors: A Systematic Review and Applicability to Neurosurgical Training	29174240	2018	Neurosurgery	Multi-Institutional	Yes
	Daly KA	Predictors for resident success in otolaryngology	16571437	2006	ENT	Single-Institution	Yes
	Burish MJ	Predicting success: What medical student measures predict resident performance in neurology?	26038279	2015	Neurology	Single-Institution	No
	Hartman ND	A Narrative Review of the Evidence Supporting Factors Used by Residency Program Directors to Select Applicants for Interviews	31210855	2019	All Subspecialties	Multi-Institutional	No
	Agarwal V	Do Residency Selection Factors Predict Radiology Resident Performance?	29239834	2018	Radiology	Single-Institution	Yes

(continued)

Table 2. Continued.

Category	1st Author	Article	PUBMED ID	Publication Year	Specialty Represented	Institutions Involved	Predictive
	Thompson RH	Predictors of Scholarly Productivity, Pursuit of Fellowship, and Academic Practice Among Urology Residents Using Medical Student Application Materials	30031833	2018	Urology	Single-Institution	Yes
	Grewal SG	Predictors of success in a urology residency program	23337683	2013	Urology	Single-Institution	Yes
	Lehman A	Predictors of a Successful Urology Resident Using Medical Student Application Materials	29432874	2018	Urology	Single-Institution	Yes
	Herndon JH	Predictors of success on the American Board of Orthopedic Surgery examination	19557490	2009	Orthopedic	Single-Institution	Yes
	Wagner JG	What Predicts Performance? A Multicenter Study Examining the Association Between Resident Performance, Rank List Position, and United States Medical Licensing Examination Step 1 Scores	27955983	2017	EM	Multi-Institutional	No
	Shellito JL	American Board of Surgery examinations: can we identify surgery residency applicants and residents who will pass the examinations on the first attempt?	19892313	2010	Gen Surg	Single-Institution	Yes
	Harffmann KL	Can performance in medical school predict performance in residency? A compilation and review of correlative studies	21612841	2011	Multi-Specialty	Multi-Institutional	No
	Chole RA	Predictors of future success in otolaryngology residency applicants	22911295	2012	ENT	Single-Institution	No
	Alterman DM	The predictive value of general surgery application data for future resident performance	22000538	2011	Gen Surg	Single-Institution	Yes
	Thompson RH	Predictors of a Successful Urology Resident Using Medical Student Application Materials	28751165	2017	Urology	Single-Institution	Yes
	Van Meter M	Does the National Resident Match Program Rank List Predict Success in Emergency Medicine Residency Programs?	27692649	2017	EM	Multi-Institutional	No
	Raman	Does Residency Selection Criteria Predict Performance in Orthopedic Surgery Residency?	25940336	2016	Orthopedic	Single-Institution	Yes
	Bhat R	Predictors of a Top Performer During Emergency Medicine Residency	26242925	2015	EM	Multi-Institutional	Yes
<i>Situational Judgment Testing:</i>							
	Gardner	Evaluation of Validity Evidence for Personality, Emotional Intelligence, and Situational	29282462	2018	Multi-Specialty	Multi-institutional	Yes

(continued)

Table 2. Continued.

Category	1st Author	Article	PUBMED ID	Publication Year	Specialty Represented	Institutions Involved	Predictive
		Judgment Tests to Identify Successful Residents					
	Gardner	Making progress on identifying those who aren't making progress: Using situational judgment tests to predict those at risk for remediation and attrition	Not on PubMed	2018	General Surgery	Single-Institution	Yes
	Strand EA	Can a structured, behavior-based interview predict future resident success?	21457921	2011	OBGYN	Single-Institution	Yes
	Ogunyemi D	Mini Surgical Simulation, Role Play, and Group and Behavioral Interviews in Resident Selection	27413446	2016	OBGYN	Single-Institution	Yes
	Burkhardt JC	Prognostic Value of the Multiple Mini-Interview For Emergency Medicine Residency Performance	25937476	2015	Emergency Medicine	Multi-institutional	Yes
<i>Standardized LORs:</i>							
	Kimple AJ	Standardized letters of recommendation and successful match into otolaryngology	26839977	2016	ENT	Single-Institution	No
	Bhat R	Predictors of a Top Performer During Emergency Medicine Residency	26242925	2015	Emergency Medicine	Multi-institutional	Yes
<i>Specialty Specific Pre-Assessment:</i>							
	Tarchala M	Pre-education Enhances the Success of Manual Training for Orthopedic Surgery Residents	30975601	2019	Orthopedics	Single-Institution	Yes
<i>Personality Traits:</i>							
	Phillips D	Personality Factors Associated With Resident Performance: Results From 12 Accreditation Council for Graduate Medical Education Accredited Orthopedic Surgery Programs	28688967	2018	Orthopedics	Single-Institution	Yes
	Merlo LJ	Personality testing may improve resident selection in anesthesiology programs	19995155	2009	Aneasthesia	Single-Institution	Yes
	Ying LD	Measuring Uncertainty Intolerance in Surgical Residents Using Standardized Assessments	31419639	2020	General Surgery	Single-Institution	Yes
	Ogunyemi DA	Associations between DISC assessment and performance in obstetrics and gynecology residents	22010523	2011	OBGYN	Single-Institution	Yes
	Quillin RC third	How residents learn predicts success in surgical residency	24209648	2013	General Surgery	Single-Institution	Yes
	Hughes BD	Personality Testing May Identify Applicants Who Will Become	30502254	2019	General Surgery	Single-Institution	Yes

(continued)

Table 2. Continued.

Category	1st Author	Article	PUBMED ID	Publication Year	Specialty Represented	Institutions Involved	Predictive
		Successful in General Surgery Residency					
	Lubelski D	Correlation of personality assessments with standard selection criteria for neurosurgical residency applicants	26848920	2016	Neurosurgery	Single-Institution	Yes

Abbreviations: ABEM, American Board of Emergency Medicine.

Comparison of Predictive Outcomes Among Evaluation Methods

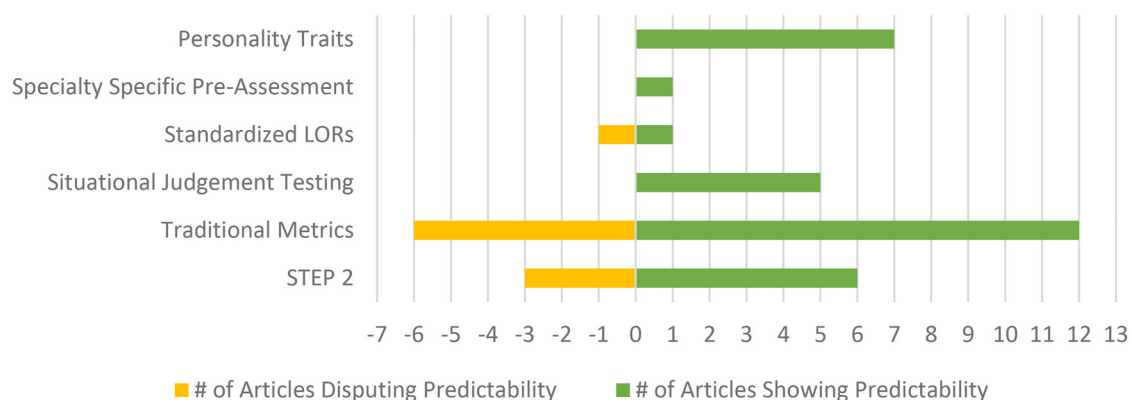


Figure 2. Comparison of the number of articles showing predictive outcomes among evaluation methods.

standardized letter of recommendation (SLOR) was a factor in determining if a resident will be successful or not.¹⁹ However, this conclusion conflicts with Kimple et al, which mentions that while the SLOR saves time and normalizes how recommendations are written, however the individuality of how people evaluate and complete the letter is suspect to inconsistency.²⁰ This discrepancy leads to a convoluted conclusion about SLORs in predicting resident success.

The goal of this scoping review was to explore innovative approaches in selecting applicants without STEP 1 scores, as such the authors found, a few papers that present tailored steps that some of the institutions have taken to add value and innovation to the resident selection process. This includes examples of institutions developing specialty specific pre-assessment testing, personality trait testing, and situational judgement testing specific to their institution to predict which students will be successful in their program. As there are several differences between the studies reviewed regarding these topics, the conclusive power of the pooled results is limited.

Specialty Specific Pre-Assessment Testing

Specialty specific pre-assessment testing was explored at a single institution for Orthopedic Surgery. The objective of

the study was to assess if pre-education of residents prior to using the high-speed drill would significantly increase their comfort level, as well as increase the success in performing a laminectomy. The study of 20 ortho residents found that residents who had “pre-education” were in fact more successful in completing the laminectomy with the high-speed drill.²¹ Pre-education in this context merely refers to education on the skill prior to assessment. This can be implemented in medical education curriculum or as an adjunct course for interested applicants. As this study doesn’t directly show that the measurement of applicant skills can provide evidence of resident success, similar technical skills like these might be useful as a pre-assessment to gauge resident performance on technical skills. Surgical specialties, especially, can use a technical assessment to predict if an applicant has the dexterity to be successful in surgery.

Situation Judgement Testing and Personality Trait Analysis

Situational judgement testing and personality trait analysis were shown to add value to the resident selection process. Situational judgement testing (SJT) has become widespread across organizations in medicine. It is used by organizations to effectively evaluate how a candidate would handle situations that

would commonly arise as an employee of that organization. In the studies included in this paper – SJT was used by institutions to effectively evaluate and compare applicants on their ability to react in specific scenarios that would likely present to a resident. This is different than using personality traits for resident selection. Personality traits are utilized in team settings to acknowledge how individuals can complement the other members of the team. Personality traits are used to better understand what individuals value and how they prefer to communicate. Organizations have found it to be helpful to know the personality traits of the people they are working with so they can create a more welcoming environment for their organization. The studies in this paper used personality traits to assess which applicants would be the best fit for their program.

Gardner et al found that Situational judgement testing could be used to incorporate more underrepresented minorities (URM). Their results show that on average, programs experienced an 8% increase in the percentage of URM candidates recommended for an interview by relying upon SJTs for interview decisions. In addition to that finding there was an increase in the URMs that were interviewed with a decrease in applicant pool size. By using USMLE cutoffs, programs can be stuck with reviewing several hundred applicants before sending out interview invitations. However, with the inclusion of SJT programs Gardner et al had an average of 55 applicants recommended for an on-site interview. They were able to decrease the number of applications to review while increasing the percentage of URMs that were invited for an interview.²²

Another study justified the cost of situational judgement testing as means to decrease the annual cost that residency programs lose on remediation intervention. Gardner et al found that administrative costs for current residency selection process range from \$45 000 to \$148 000 per year per program. They found that about 30% of residents require at least 1 remediation intervention for concerns around non-technical skills, costing programs \$3400 to \$5300 per episode. Attrition rates ranged from about 20% to 40% for the general surgery residencies in the study.²³ Although it may cost programs to implement additional interviews, having an extensive way to predict which residents will not require remediation could save programs money. Lastly, articles in the personality trait category were able to provide insight of the institution in the study's selection process.

Bell et al found that utilizing a third-party personality trait assessment with other traditional criteria could be useful in selecting the residents that would most easily integrate into program's culture. They found that the personality trait assessment was useful in (1) reassuring that the decisions made in the ranking list were compatible with the program, (2) identifying which specific candidates would fit best, (3) providing assurance that selected applicants would not conflict with expectations, and (4) identifying which applicants would need additional coaching versus the applicants that would be self-directed

learners. In addition, the personality trait assessment alerted the program how to best communicate, motivate, and incorporate constructive criticism to its future residents.²⁴

Beyond the 5 categories of methods of resident selection mentioned in the results, there were a few other noteworthy methods that were identified during the review. These studies did not have any predictability metrics, so they were excluded from the 5 main categories evaluated in this study. Some articles explored ways to limit the number of applications – hoping to allow for more time to review applications. In order to decrease the size of application pools numbers, authors suggested to increase the cost of applications, adding supplemental applications, and having an early application cycle for top choices.² However, these propositions disproportionately would affect some of our more underrepresented applicants. In addition, these articles not having a component that compared how limiting the number of applications can increase the ability to predict resident success those articles were excluded.

Two other innovative ideas for resident selection that were found in the literature review, include utilizing a Standard Video Interview (SVI)²⁵ and a proposed “STAR” model in the narrowing of the application pool.²⁶

The Association of American Medical Colleges (AAMC) with Accreditation Council of Graduate Medical Education (ACGME)-accredited emergency medicine programs launched a pilot program for the 2018 residency application cycle that aimed to measure interpersonal and communication skills. The pilot program was a standardized video interview (SVI) that applicants completed in conjunction to their application. The interviews were scored by a third party trained by the AAMC and given to emergency medicine program directors. Radabaugh et al concluded that these SVIs could broaden the net that our current application cutoffs are missing and create a more diverse applicant pool.²⁷ Bird et al found that SVI total scores correlated in expected ways with other selection data.²⁵ There were non-significant correlations between SVI total scores and theoretically unrelated academic variables (ie, Step 1 scores and AOA membership). This finding is encouraging because it suggests that SVI scores measure something other than academic performance and may add valuable information to the selection process. Possibly using the two metrics of the SVI and Electronic Residency Application Service (ERAS) in conjunction could be useful for program directors when finalizing a rank list.

Another innovative method to decrease the amount of labor associated with identifying the best applicants for a program to interview is using algorithms to automate the selection process. Villwock et al conducted a retrospective cohort study at a single institution to evaluate the effectiveness of a such tool.²⁶ They designed a Selection Tool for Applicants to Residency (STAR) as an objective approach to select students for an interview without a USMLE cutoff. They found that the applicants

selected by the STAR correlated to applicants with equal or better USMLE scores. In addition, there was no significant difference in the proportion of women or underrepresented minorities invited without STAR. They concluded that the STAR algorithm significantly shortened the time programs need to evaluate an applicant pool without affecting the composition of interview groups. Evaluation systems like the STAR method can possibly shape the future of application evaluations. Although, the STAR method is able to decrease the amount of time needed to develop a pre-interview rank list, it may not be an innovation in residency selection. When comparing the STAR method to a pre-STAR selection method, the demographics of the pool were similar. Lastly, there is no evidence that STAR can predict which applicants will be most successful in their program. Rather, the STAR method proposes a more streamlined path to get to the same destination.

As with all scoping reviews, a limitation in this scoping review was the heterogeneity of the selected literature. The data from the included articles in this review were highly variable in the way they were presented as well as methods of study, which made it cumbersome to effectively compare and draw conclusions of the residency application factors that could be predictable of resident success. Some articles relied heavily on the use of specialty board exam pass rates as an indicator of resident success. Some of the other categories evaluated an applicant's more holistic qualities and relied on residency faculty evaluations, communication skills, behavioral tendencies, and selection for chief resident/awards as a measurement of resident success. Some articles were unconventional as they evaluated the predictability of whether a resident would not have to remediate as an indicator of resident success. Some articles were more specific. For example, the only article in the Specialty Specific Pre-Assessment category found that orthopedic residents who had pre-education/and an assessment were in fact more successful in completing the laminectomy with the high-speed drill – obviously this article only used a specific procedure as a measurement for resident success. These mixed definitions of resident success make it difficult to compare the categories against each other. However, does allow some opportunity to judge the effectiveness of the individual factor in predicting resident success within its own category. Another limitation of this review is that most of the articles included in this study were single institutional, single specialty samples. This plays an important role as results can vary by how program directors evaluate applicants, the competitiveness among the applications, and the institution's standards. Students applying into more competitive specialties are more likely to have similar scores and therefore be compared within a narrow margin of honors in medical school, number of publications, awards, standardized test scores. Within these specialties, program directors have a varying subset of characteristics they use for selection. The articles become more difficult to compare when considering that an institution's specialty

program vary drastically from another institution's specialty program. For the studies that evaluated personality traits as a factor that determined resident success in a single subspecialty, the personality traits that showed predictiveness in a specific specialty varied among institutions. Therefore, each subspecialty program director shouldn't generalize certain personality traits and their respective subspecialty – they would need to consider the environment of their individual program to predict which resident personalities will be more likely to complement each other. A final limitation of the study is the different number of studies available to evaluate in each of the 5 categories was non consistent. Although these limitations apply to the inability to perform a pooled data analysis, the objective of this study was to identify and present the studies available showing innovative methods of resident selection, which the authors believe this study accomplishes.

Conclusion

USMLE Step 1's change from numerical score reporting to pass/fail grading naturally compels Resident Selection Committees (RSCs) to explore alternative, innovative methods through which they can evaluate and pinpoint "optimal" applicants. This scoping review explored the literature to identify which methods of resident selection are useful in finding the applicant that will be the most successful resident. Many programs utilize STEP 1 scores to stratify applicants however, this has apparently over-emphasized an exam that was not designed to differentiate students while negatively affecting underrepresented minorities in medicine. Here, the authors found personality traits, situational judgement testing, and specialty specific assessments to be predictive in selecting successful residents. These methods provided some added value that the traditional application process hadn't. Methods that were inconsistent in selecting successful residents were Step 2 scores, Traditionally Used Metrics [core clerkship scores, interview performance, musical talent, sports involvement, AOA membership, research publications, school unprofessional behavior, Dean's letters, Rank list], and Standardized Letters of Recommendation. Further research should aim to analyze exactly how RSCs utilize these assessment tools to aid in screening their large and competitive applicant pools to find residents that will be successful in their program.

Ethical Approval

Not applicable, because this article does not contain any studies with human or animal subjects.

Informed Consent

Not applicable, because this article does not contain any studies with human or animal subjects.

Trial Registration

Not applicable, because this article does not contain any clinical trials.

REFERENCES

- Katsufakis PJ, Chaudhry HJ. Improving residency selection requires close study and better understanding of stakeholder needs. *Acad Med.* 2019;94(3):305-308. doi:10.1097/ACM.0000000000002559
- Katsufakis PJ, Uhler TA, Jones LD. The residency application process: pursuing improved outcomes through better understanding of the issues. *Acad Med.* 2016;91(11):1483-1487. doi:10.1097/ACM.0000000000001411
- de Haan JB, Markham T, Ghebremichael S. Considerations of a resident recruitment committee on the USMLE step 1 examination. *Acad Med.* 2019;94(7):922-923. doi:10.1097/ACM.0000000000002721
- Rozenshtein A, Mullins ME, Marx MV. The USMLE step 1 pass/fail reporting proposal: the APDR position. *Acad Radiol.* 2019;26(10):1400-1402. doi:10.1016/j.acra.2019.06.004
- Deng F, Gluckstein JA, Larsen DP. Student-directed retrieval practice is a predictor of medical licensing examination performance. *Perspect Med Educ.* 2015;4(6):308-313. doi:10.1007/s40037-015-0220-x
- Swanson DB, Sawhill A, Holtzman KZ, et al. Relationship between performance on part I of the American board of orthopaedic surgery certifying examination and scores on USMLE steps 1 and 2. *Acad Med.* 2009;84(10 Suppl):S21-S24. doi:10.1097/ACM.0b013e3181b37fd2
- Dougherty PJ, Walter N, Schilling P, Najibi S, Herkowitz H. Do scores of the USMLE step 1 and OITE correlate with the ABOS part I certifying examination?: a multicenter study. *Clin Orthop Relat Res.* 2010;468(10):2797-2802. doi:10.1007/s11999-010-1327-3
- Dyrstad BW, Pope D, Milbrandt JC, Beck RT, Weinhoef AL, Idusuyi OB. Predictive measures of a resident's performance on written orthopaedic board scores. *Iowa Orthop J.* 2011;31:238-243.
- McGaghie WC, Cohen ER, Wayne DB. Are United States medical licensing exam step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? *Acad Med.* 2011;86(1):48-52. doi:10.1097/ACM.0b013e31811facdb
- Reed DA, Shanafelt TD, Satele DW, et al. Relationship of pass/fail grading and curriculum structure with well-being among preclinical medical students: a multi-institutional study. *Acad Med.* 2011;86(11):1367-1373. doi:10.1097/ACM.0b013e3182305d81
- Bloodgood RA, Short JG, Jackson JM, Martindale JR. A change to pass/fail grading in the first two years at one medical school results in improved psychological well-being. *Acad Med.* 2009;84(5):655-662. doi:10.1097/ACM.0b013e31819f6d78
- Rubright JD, Jodoin M, Barone MA. Examining demographics, prior academic performance, and United States medical licensing examination scores. *Acad Med.* 2019;94(3):364-370. doi:10.1097/ACM.0000000000002366
- Raman T, Alrabaa RG, Sood A, Maloof P, Benevenia J, Berberian W. Does residency selection criteria predict performance in orthopaedic surgery residency? *Clin Orthop Relat Res.* 2016;474(4):908-914. doi:10.1007/s11999-015-4317-7
- Burkhardt JC, Parekh KP, Gallahue FE, et al. A critical disconnect: residency selection factors lack correlation With intern performance. *J Grad Med Educ.* 2020;12(6):696-704. doi:10.4300/JGME-D-20-00013.1
- Hu AC, Gu JT, Wong BJF. Objective measures and the standardized letter of recommendation in the otolaryngology residency match. *Laryngoscope.* 2020;130(3):603-608. doi:10.1002/lary.28054
- Haist SA, Butler AP, Paniagua MA. Testing and evaluation: the present and future of the assessment of medical professionals. *Adv Physiol Educ.* 2017;41(1):149-153. doi:10.1152/advan.00001.2017
- Brandenburg S, Kruzick T, Lin CT, Robinson A, Adams LJ. Residency selection criteria: what medical students perceive as important. *Med Educ Online.* 2005;10(1):4383. doi:10.3402/meo.v10i.4383
- Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.* 2018;169(7):467-473. doi:10.7326/M18-0850
- Bhat R, Takenaka K, Levine B, et al. Predictors of a Top performer during emergency medicine residency. *J Emerg Med.* 2015;49(4):505-512. doi:10.1016/j.jemermed.2015.05.035
- Kimple AJ, McClurg SW, Del Signore AG, Tomoum MO, Lin F-C, Senior BA. Standardized letters of recommendation and successful match into otolaryngology. *Laryngoscope.* 2016;126(5):1071-1076. doi:10.1002/lary.25637
- Tarchala M, Charbonneau M, Aduljabbar FH, Teles AR, Weber M. Pre-education enhances the success of manual training for orthopedic surgery residents. *J Surg Educ.* 2019;76(5):1433-1439. doi:10.1016/j.jsurg.2019.03.005
- Gardner AK, Cavanaugh KJ, Willis RE, Dunkin BJ. Can better selection tools help Us achieve Our diversity goals in postgraduate medical education? Comparing Use of USMLE step 1 scores and situational judgment tests at 7 surgical residencies. *Acad Med.* 2020;95(5):751-757. doi:10.1097/ACM.0000000000003092
- Gardner AK, Grantcharov T, Dunkin BJ. The science of selection: using best practices From industry to improve success in surgery training. *J Surg Educ.* 2018;75(2):278-285. doi:10.1016/j.jsurg.2017.07.010
- Bell RM, Fann SA, Morrison JE, Lisk JR. Determining personal talents and behavioral styles of applicants to surgical training: a new look at an old problem, part I. *J Surg Educ.* 2011;68(6):534-541. doi:10.1016/j.jsurg.2011.05.016
- Bird SB, Hern HG, Blomkalns A, et al. Innovation in residency selection: the AAMC standardized video interview. *Acad Med.* 2019;94(10):1489-1497. doi:10.1097/ACM.0000000000002705
- Villwock JA, Hamill CS, Sale KA, Sykes KJ. Beyond the USMLE: the STAR algorithm for initial residency applicant screening and interview selection. *J Surg Res.* 2019;235:447-452. doi:10.1016/j.jss.2018.07.057
- Radabaugh CL, Hawkins RE, Welcher CM, et al. Beyond the United States medical licensing examination score: assessing competence for entering residency. *Acad Med.* 2019;94(7):983-989. doi:10.1097/ACM.0000000000002728