



Correlation of Ophthalmology Residency Application Characteristics with Subsequent Performance in Residency

Brett M. Gudgel, MD¹ Andrew T. Melson, MD¹ Justin Dvorak, PhD¹ Kai Ding, PhD¹
R. Michael Siatkowski, MD¹

¹University of Oklahoma Health Science Center, Oklahoma City, Oklahoma

Address for correspondence Andrew T. Melson, MD, Dean McGee Eye Institute, 608 Stanton L. Young Boulevard, Oklahoma City, OK 73104 (e-mail: andrew-melson@dmei.org).

J Acad Ophthalmol 2021;13:e151–e157.

Abstract

Purpose Only from reviewing applications, it is difficult to identify which applicants will be successful ophthalmology residents. The change of the USMLE Step 1 scoring to “Pass/Fail” removes another quantitative metric. We aimed to identify application attributes correlated with successful residency performance. This study also used artificial intelligence (AI) to evaluate letters of recommendation (LOR), the Dean’s letter (MSPE), and personal statement (PS).

Design Retrospective analysis of application characteristics versus residency performance was conducted.

Participants Residents who graduated from the Dean McGee Eye Institute/University of Oklahoma Ophthalmology residency from 2004 to 2019 were included in this study.

Methods Thirty-four attributes were recorded from each application. Residents were subjectively ranked into tertiles and top and bottom deciles based on residency performance by faculty present during their training. The Ophthalmic Knowledge Assessment Program (OKAP) examination scores were used as an objective performance metric. Analysis was performed to identify associations between application attributes and tertile/decile ranking. Additional analysis used AI and natural language processing to evaluate applicant LORs, MSPE, and PS.

Main Outcome Measures Characteristics from residency applications that correlate with resident performance were the primary outcome of this study.

Results Fifty-five residents and 21 faculty members were included. A grade of “A” or “Honors” in the obstetrics/gynecology (OB/GYN) clerkship and the presence of a home ophthalmology department were associated with ranking in the top tertile but not the top decile. Mean core clerkship grades, medical school ranking in the top 25 U.S. News and World Report (USNWR) primary care rankings, and postgraduate year (PGY)-2 and PGY-3 OKAP scores were predictive of being ranked in both the top tertile and the top decile. USMLE scores, alpha-omega-alpha (AOA) status, and number of publications did

Keywords

- ▶ Residency Application
- ▶ resident performance
- ▶ artificial intelligence

received
November 28, 2020
accepted after revision
April 3, 2021

DOI <https://doi.org/10.1055/s-0041-1733932>.
ISSN 2475-4757.

© 2021. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Thieme Medical Publishers, Inc., 333 Seventh Avenue, 18th Floor, New York, NY 10001, USA

not correlate with subjective resident performance. AI analysis of LORs, MSPE, and PS did not identify any text features that correlated with resident performance.

Conclusions Many metrics traditionally felt to be predictive of residency success (USMLE scores, AOA status, and research) did not predict resident success in our study. We did confirm the importance of core clerkship grades and medical school ranking. Objective measures of success such as PGY-2 and PGY-3 OKAP scores were associated with high subjective ranking.

Ophthalmology is one of the most competitive residencies within medicine.^{1,2} As such, the number and competitiveness of applicants continue to grow.¹ In 2019, for the 790 medical students who initially registered for the ophthalmology match, 649 applicants submitted an application. Of these applicants, 25% did not match into an ophthalmology residency program.¹ Residency programs are tasked with sorting through hundreds of applications to fill a mean of four residency slots per year. In an effort to better screen the large number of applicants, studies have attempted to identify which selection factors are most valuable in screening for top-tier candidates. Factors, such as interview performance, clinical course grades, recommendation letters, and USMLE scores have all been determined to be among the most important factors for applicant selection.^{3,4} Though these factors have been identified as important among interview selection committees, no studies have been done within ophthalmology to show if these attributes have any predictive value in applicant performance throughout residency. This uncertainty is highlighted by the fact that fewer than 50% of respondents felt that there was a strong correlation between the highest ranked applicants and the best performers during residency.³ To further complicate the situation, the United States Medical Licensing Examination (USMLE) program has elected to change its score reporting for the step-1 examination from a three-digit numeric score to a pass/fail outcome, eliminating one of the most commonly used application screening metrics.⁵ A clear need for long-term studies correlating applicant attributes with successful residency performance has been identified.³

Multiple studies have been conducted within other specialties to attempt to identify the attributes that best correlate with resident performance.⁶⁻³¹ Many of these studies fail to agree on common predictive resident attributes, and often contradict each other. There is significant variability in predictive factors from specialty to specialty which further highlights the need for the specialty of ophthalmology to attempt to identify ophthalmology-specific predictors of success. Better determining such predictors could help programs identify those applicants who are most likely to become excellent residents. Additional difficulty exists with the interpretation and utilization of letters of recommendation (LOR) and the Dean's letter (MSPE) in the resident applicant evaluation process. Program directors and resident selection committees have indicated that LOR are among the most important factors in resident selection.³ Despite this,

there remains inconsistency regarding interviewer reliability, and there is some doubt that LOR are actually useful in predicting future applicant success as a resident.³² The same holds true with respect to the personal statement (PS), with no identifiable studies correlating its contents to success in residency.

The goals of this study are to (1) identify resident attributes that are predictive of success in an ophthalmology residency, (2) validate or invalidate the selection metrics that are currently most highly valued, and (3) attempt to utilize artificial intelligence (AI) and machine learning to develop a more meaningful and consistent method of LOR, MSPE, and PS interpretation and utilization. In doing so, we hope to improve the resident selection process within ophthalmology by giving programs additional evidence to guide application evaluation.

Materials and Methods

Application Characteristics Collected

This study was granted exempt status by the Institutional Review Board of the University of Oklahoma Health Sciences Center. The residency applications of residents graduating from the Dean McGee Eye Institute/University of Oklahoma Ophthalmology Department from 2004 through 2019 (55 residents) were obtained through the San Francisco Match database and Dean McGee Eye Institute records. The applications were reviewed, and the applicant characteristics that were recorded included: gender, number of preresidency publications, USMLE step-1 score, USMLE step-2 score (if listed), alpha-omega-alpha (AOA) status, number of "away" or "audition" rotations, advanced degrees (e.g., PhD, MS, and others), preclinical medical school grade point average (GPA), core clerkship grades, ophthalmology rotation grade, military service, previous nonmedical employment, participation in a varsity sport in college, significant experience in art or music, undergraduate GPA, ranking of medical school (based on 2019 U.S. News and World Report [USNWR] rankings), ranking of undergraduate school (2019 USNWR ranking), majoring in a science; presence of a home ophthalmology department, ranking of ophthalmology department (based on 2019 Doximity rankings), attendance of a medical school in a nearby state, and attendance of an undergraduate school in a surrounding state. The clerkships of medicine, surgery, pediatrics, obstetrics/gynecology (OB/GYN), and psychiatry were considered to be "core." The highest possible

grade (A, Honors, etc.) was considered to be Honors for the rotation. Both the USNWR primary care and research rankings were used for medical school ranking. Significant experience in art or music was defined as consistent involvement for at least a decade, a history of multiple public performances/exhibitions/competitions, or at least one prize, such as a state, regional, or national level.

Resident Performance Determination

Resident performance was composed of both a subjective and an objective component. For the subjective component, a survey was built on the RedCap platform and was sent to every current faculty member at the Dean McGee Eye Institute/University who was also a member of the faculty during the review period. The survey asked, "In which tertile would you rank the above resident as compared to your expectation for a graduating resident at completion of training?" If the top tertile was selected, the survey asked, "Is the resident in the top 10%?" and if the bottom tertile was selected, the survey asked, "Is the resident in the bottom 10%?" The survey then assessed each resident for deficiencies in each of the Accreditation Council for Graduate Medical Education (ACGME) core competencies both during residency and at the completion of residency. Space was provided for additional comments. The faculty members were blinded to the resident applications. At our institution, the faculty members are also blinded to the resident Ophthalmic Knowledge Assessment Program (OKAP) scores. Resident scores on postgraduate year (PGY)-2, PGY-3, and PGY-4 OKAP examinations were used as an objective measurement of residency performance, and had no relationship to the subjective faculty ranking.

Statistical Analysis

Data were descriptively summarized (e.g., mean, SD, range, count, and percentage). The effect of each predictor on the resident's ranking (top vs. bottom tertiles and top 10% vs. bottom 90%) was assessed with a logit model with the compound-symmetry covariance structure, accounting for multiple ratings per resident.

Interrater reliability was assessed using Fleiss' kappa statistic for multiple categories and raters, which was computed according to Chen et al.³²

The effect of predictors on numerical OKAP scores was investigated using a series of linear regression models, as each resident had only a single OKAP score for each clinical year.

To construct a single resident-quality metric for the purposes of natural language processing (NLP), the individual raters' assignments of tertiles and top versus bottom 10% were transformed into numerical percentile ratings, and then aggregated within residents. The transformations were designed such that the resulting numerical value represented the midpoint of the "bin" in which a given rater would place a resident. Note that top and bottom tertile assignments are further refined by top 10% and bottom 10% status, respectively; for example, a resident who was placed in the top tertile by a given rater, but not the top 10%, would

receive a score of 78.3% from that rater because this is the midpoint between 66.7% (lower end of top tertile) and 90% (upper end of top tertile after excluding the top 10%).

For the AI component, we performed a token analysis and sentiment analysis to determine whether any text features were predictive of resident performance.

Analyses were performed on three document classes: MSPE, PS, and LOR. Token analysis was conducted by stemming all words in each document class and applying term-frequency and inverse-document frequency weighting. Sentiment analysis was conducted by mapping each word or token in a given document to one of four NLP corpora to generate a set of scores for each.³³⁻³⁶ Sentiment scores in each of the corpora were transformed to a numeric range, and harmonized for modeling purposes. Following this process, each sentiment score for the document was used in two different linear regression models as a predictor for each of the associated resident's outcomes (percentiles and OKAP scores). All statistical analyses, including NLP modeling, were conducted in R v3.5.1 and SAS v9.4, with an α level of 0.05 (two-sided) defining statistical significance.

Results

A total of 55 residents from 2004 to 2019 were included in the study. A summary of resident characteristics can be found in ►Table 1. The majority of residents in the study period were male (80%). The mean USMLE step-1 score was 243.42 (range: 203-270), and the mean USMLE step-2 score was 246.50 (range: 228 - 273, though 43 applications did not report a step-2 score). The majority of residents were elected to AOA (61%, with 11 applications not reporting any kind of AOA status). The mean number of preresidency publications

Table 1 Baseline characteristics and attributes of residents evaluated

Resident characteristics	
Total number of residents	55
Male	44 (80.00%)
Female	11 (20.00%)
Mean USMLE step 1	243.42 ± 13.31
Mean USMLE step 2	246.50 ± 14.61
AOA membership ^a	34 (61.82%)
Mean pre-residency publications	1.04 (range: 0-8)
Advanced degree	3 (5.45%)
Military service	2 (3.64%)
College varsity athlete	3 (5.45%)
Significant art/music experience	9 (16.36%)
Home ophthalmology program	49 (89.09%)
Home ophthalmology program ranked within top 25	20 (36.36%)

Abbreviations: AOA, alpha-omega-alpha Honor Medical Society; USMLE, United States Medical Licensing Examination.

^a11 applications did not report any AOA status.

Table 2 Resident attributes predictive of tertile rating on faculty survey

Predictor	Top tertile	Bottom tertiles	p-Value
	(n = 406)	(n = 395)	
Honors in OB/GYN clerkship	307 (75.99%)	233 (60.05%)	0.0400
Presence of home ophthalmology department	382 (94.09%)	337 (85.32%)	0.0356
Medical school ranked in USNWR top 25 (primary care) ^a	67 (16.50%)	31 (7.85%)	0.0350
Mean core clerkship GPA ^a	3.91 ± 0.17	3.79 ± 0.28	0.0248
Mean PGY-2 OKAP score (percentile) ^a	67.19 ± 24.02	56.94 ± 28.22	0.0214
Mean PGY-3 OKAP score (percentile) ^a	75.26 ± 18.68	68.32 ± 20.59	0.0133
Undergraduate school ranked in USNWR top 50 ^b	69 (17.47%)	117 (30.87%)	0.0371

Abbreviations: GPA, grade point average; OB/GYN, obstetrics and gynecology; OKAP, ophthalmic knowledge assessment program; PGY, postgraduate year; USNWR, United States News and World Report.

^aAlso predictive of ranking in top decile.

^bInversely correlated with ranking in top tertile and decile.

was 1, ranging from 0 to 8. Only three residents had advanced degrees, two residents served in the military, and three residents played a varsity sport in college. The majority of applicants had a home ophthalmology department (89%), with 36% of home ophthalmology departments ranking in the top 25 based on the 2019 Doximity rankings.

Twenty-one faculty members participated in the survey. The mean number of reviewers per resident was 14.6 (range: 9–21). Interreviewer reliability for tertile ranking, measured by Fleiss' kappa, was 0.148 ($p < 0.001$), which suggests only a slight agreement among the reviewers. Predictors of subjective success based on the faculty survey can be found in **Table 2**. The analysis was done comparing residents in the top tertile to the bottom tertiles, as well as the top 10% to the bottom 90%. Applicant characteristics that were predictive of being ranked in the top tertile were a grade of A or Honors in OB/GYN clerkship ($p = 0.040$), presence of a home ophthalmology department ($p = 0.036$), medical school ranking in the top 25 USNWR primary care rankings ($p = 0.035$), and mean core clerkship GPA ($p = 0.025$). Characteristics that were predictive of being in the top 10% were attending a medical school in a surrounding state ($p = 0.047$), medical school ranked in the top 25 USNWR primary care rankings ($p = 0.022$), and mean core clerkship GPA ($p = 0.010$). Attending an undergraduate school ranking in the USNWR top 50 was inversely correlated with being ranked in both the top tertile and top 10% ($p = 0.037$ and 0.027 , respectively). Performance on the PGY-2 ($p = 0.021$ and 0.006 , respectively) and PGY-3 ($p = 0.013$ and $p = 0.004$, respectively) OKAP examinations was predictive of being subjectively ranked in both the top tertile and top 10%.

With respect to objective success as measured by OKAP examinations, USMLE step-1 score was predictive of performance in the PGY-2 (0.9 point increase in OKAP score per 1 point increase in USMLE step-1 score, $p = 0.001$), PGY-3 (0.79 point increase in OKAP score per 1 point increase in USMLE step-1 score, $p < 0.0001$), and PGY-4 years (0.51 point increase in OKAP score per 1 point increase in USMLE step-1 score, $p = 0.014$). Undergraduate GPA ($p = 0.020$) and attending a medical school ranked in USNWR top 25 primary care

ranking ($p = 0.006$) were predictive of performance on PGY-2 OKAP examination. Male gender ($p = 0.007$), preclinical medical school GPA ($p = 0.049$), and Honors in surgery clerkship ($p = 0.016$) were predictive of performance on PGY-3 OKAP examination. Significant experience in art or music was inversely predictive of performance on both the PGY-2 ($p = 0.028$) and PGY-3 ($p = 0.015$) OKAP examinations. Characteristics that were notably not predictive of resident success based on the faculty survey and OKAP results included membership in AOA, number of publications, average number of nonmedical employments, and grade of Honors or A in the medicine clerkship. USMLE step-1 and -2 scores were not predictive of subjective resident success based on faculty survey but did predict OKAP performance in PGY-2/-3 years.

In the NLP analyses, we were unable to identify any words or phrases in any of the documents that correlated nonrandomly with either resident performance or OKAP scores. In the token analysis, only the MSPE had reasonable out-of-bag accuracy (random forest) and leave-one-out cross-validation accuracy; however, this high accuracy was primarily an artifact of the ranking distribution and resultant confusion matrix, rather than indicative of predictive keywords. In the sentiment analysis, only a limited number of sentiments were associated with the overall resident performance score, and the consistency of this association did not hold across document classes. No sentiment scores were consistently associated with OKAP scores across residency years.

Discussion

To our knowledge, this represents the first published, large-scale, multifaculty review study within ophthalmology that correlates preridency characteristics with residency performance. Additionally, this study is the first to evaluate the use of AI as a potential predictor of which applicants will be successful in residency. Prior studies have utilized machine learning to evaluate LORs for gender bias,^{37–41} racial bias,⁴² and even success rates for matching into residency,⁴³ but none have evaluated for success during residency. Our study

identified multiple factors that were predictive of subjective and objective resident success, some of which differed from studies in other subspecialties.

Our study used both a subjective and an objective component to define success in residency. The following attributes were predictive of subjective success as measured by our faculty survey: mean core clerkship GPA (top tertile and top decile), attending a medical school ranked in the USNWR top 25 primary care ranking (top tertile and top decile), attending a medical school in a surrounding state (top decile only), presence of a home ophthalmology program (top tertile only), and grade of Honors or "A" in the OB/GYN clerkship (top decile only). Performance on PGY-2 and PGY-3 OKAP examinations was also predictive of top tertile and top decile subjective rating, but attributes prior to residency training are not known. For the objective component of resident success, our study used performance on the OKAP examination as a metric. The only application attribute that was predictive of OKAP performance on all 3 years was the USMLE step-1 score.

There has been some concern over the usefulness of clerkship grades as an evaluation metric which is highlighted by the notable interapplicant variability in methods of clerkship grade reporting.^{44,45} Despite these concerns, performance on clinical clerkships appears to be one of the most consistently reported predictors of both resident clinical and academic performance.⁶⁻¹⁶ This held true in our study as well, with core clerkship performance being one of the most significant predictors of subjective resident performance, even though the mean differences were rather small. Logically, this makes sense, as residents who were able to master the education, patient care, and social aspects of clinical rotations are likely to show continued success in future clinical scenarios during residency. Similar logic applies to the predictive value of an applicant training in a top 25 primary-care medical school, in that the work ethics and academic achievement required to attend these medical schools, as well as the quality of training at such institutions, would likely produce skills that enable strong clinical performance as a resident.

Though not as frequently reported, several other studies have highlighted the predictive nature of medical school ranking for future resident performance.¹⁶⁻²⁰ Of particular interest in our study was the predictive value of the OB/GYN clerkship in top tertile residents, the only individual core clerkship with such correlation. Two other studies have also shown that a grade of Honors in the OB/GYN clerkship can be predictive of resident success.^{6,7} There are several possible explanations for this finding as follows: (1) the unique clinical and social scenarios that occur in the delivery of OB/GYN care require a special level of social and communication skills; (2) most medical students realize that exceptional performance on the "major" clerkships (i.e., internal medicine and surgery) is required for a competitive application, but the students who are less inclined to give 100% effort at all times may not put forth the effort needed to Honors a clerkship as demanding as OB/GYN; and (3) OB/GYN faculty and residents may grade harder than other

specialties. The predictive value of this metric reinforces the importance of reporting clinical clerkship grades. As some schools transition to pass/fail core clerkship grading, their students could be placed at a relative competitive disadvantage in the residency selection process.

Some commonly valued application attributes that were not found to be predictive of subjective resident performance included USMLE step-1 and -2 scores, membership in AOA, number of nonmedical employments, and number of preresidency publications. Our study also evaluated the correlation between resident performance on the OKAP in-service examination and subjective faculty ranking. At our institution, the only faculty members, who are aware of the numerical resident OKAP scores, are the Department Chair and the Program Director. Performance on the PGY-2 and PGY-3 OKAP examination was predictive of being ranked in both the top tertile and the top 10%. Since the faculty are blinded to resident OKAP scores, our data appear to validate the positive correlation of the PGY-2 and PGY-3 OKAP examinations with resident clinical performance, as well as the usefulness of the OKAP examination as a metric for resident success. Performance on the PGY-4 OKAP examination was not found to be predictive which is likely due to improvement in the OKAP scores of the bottom tertiles, as knowledge gaps were addressed and improved throughout residency.

While USMLE step-1 score was the only application attribute predictive of objective success as measured by OKAP performance across all 3 years, the degree to which a higher USMLE step-1 score correlated with a higher OKAP score decreased throughout residency. This may be explained by a decreased reliance on raw test-taking skills and an increased reliance on learned ophthalmic knowledge acquired throughout residency. Interestingly, significant experience in art or music was found to be negatively predictive of performance on the PGY-2 and PGY-3 OKAP examination. This finding may be an example of the divergent and convergent nature of creative problem solving in medicine versus the arts.

The AI component of our study yielded no useful correlations. This likely results from the nonuniform writing styles of LOR and PS authors, with many from different specialties, at different levels of their career, and with different innate linguistic skills. Similarly, the MSPE style and format is very different at different institutions. Finally, there is likely an inherent attempt to put forth the best impression for each applicant, and relative reticence to note negatives unless they are egregious.

Limitations and Strengths

Our study has several limitations. One of the biggest difficulties common to studies of this kind is defining what actually constitutes a "successful" resident. Like many studies, we divided this definition into a subjective and objective component. While we would consider the number of faculty members captured in our subjective component survey as a strength, not every resident was evaluated by the same number of faculty members. Though our survey showed statistical significance and interrater reliability, suggesting

only slight agreement among reviewers. Some faculty members may be harder graders than others, creating a nonstandard evaluation for the residents.

Ophthalmology is a relatively small specialty with relatively small residency classes. Because of this, we are unable to include as many residents in our study as larger specialties such as internal medicine or general surgery. This limits our ability to draw conclusions about less common applicant characteristics such as varsity sports participation ($n = 3$). Additionally, ophthalmology is a very competitive specialty, and our residency typically receives 100 applications per slot available. Thus, the applicants who are interviewed have already been prescreened based on several application metrics. In light of this, our results are likely not generalizable to the entire, nonscreened ophthalmology applicant pool. Additionally, the subjective USNWR and Doximity rankings change over time, and the use of only the most recent year's rankings does not take into account these changes. Furthermore, our study population was 80% male during the time period studied. Research in other specialties have demonstrated gender-based differences in LORs.^{37,42,43} Though gender was only found to be a significant predictor for PGY-3 OKAP scores in our study, gender differences between our study population and the general applicant pool serves as a potential source of bias.

Finally, this study is limited by the inability to understand the variation in applicant characteristics present among an entire application cohort. We obviously could not examine the documents of all applicants to our program, including applicants who were screened out early in the process without an interview, applicants who were interviewed but not offered a position, and applicants who were offered a position but matched elsewhere. It is certainly possible that a study of this nature, although virtually impossible to perform, may identify many more significant differences among these groups of candidates.

Conclusion

In summary, to our knowledge this represents the first large-scale, multifaculty review study within ophthalmology that correlates residency application characteristics with actual residency performance. To our knowledge, this is also the first study within any specialty to use machine learning and AI to analyze the quality of LOR, MSPE, and PSs with regard to their predictive nature for success in that specialty. While core clerkship performance (especially in OB/GYN) and medical school ranking correlated with future success as a resident in our program, many commonly valued application characteristics, such as USMLE scores, AOA status, and number of preresidency publications, were not predictive of subjective success in residency. USMLE step-1 scores were decreasingly predictive over time of OKAP performance throughout residency. Our inability to identify predictive language in the current LOR formatting suggests that this aspect of the application process may be less discerning. While a standardized LOR process has been proposed, it has not been broadly adopted and faces barriers to widespread

implementation. Finally, our study lends credence to the use of the OKAP examination as a surrogate of more global resident performance.

Meeting Presentation

This study was previously presented in part as a poster at the annual Association of University Professors of Ophthalmology (AUPO) meeting, 2019.

Conflict of Interest

No conflicting relationship exists for any author.

Acknowledgments

This research is supported in part by an unrestricted grant from Research to Prevent Blindness, Inc., New York City, NY. The sponsor or funding organization had no role in the design or conduct of this research.

References

- 1 Ophthalmology Match Summary Report 2019. Accessed May 7, 2020 at: https://www.sfmach.org/PDFFilesDisplay/Ophthalmology_Residency_Stats_2019.pdf
- 2 National Resident Matching Program. Results and Data: 2018 main residency match. Accessed May 7, 2020 at: <http://www.nrmp.org/wp-content/uploads/2018/04/Main-Match-Result-and-Data-2018.pdf>
- 3 Nallasamy S, Uhler T, Nallasamy N, Tapino PJ, Volpe NJ. Ophthalmology resident selection: current trends in selection criteria and improving the process. *Ophthalmology* 2010;117(05):1041–1047
- 4 Siatkowski RM, Mian SI, Culican SM, et al; Association of University Professors of Ophthalmology. Probability of success in the ophthalmology residency match: three-year outcomes analysis of San Francisco matching program data. *J Acad Ophthalmol* 2018; 10(01):e150–e157
- 5 United States Medical Licensing Examination. Change to pass/fail score reporting for step 1. Accessed May 7, 2020 at: <https://www.usmle.org/incus/>
- 6 Thompson RH, Lohse CM, Husmann DA, Leibovich BC, Gettman MT. Predictors of a successful urology resident using medical student application materials. *Urology* 2017;108:22–28
- 7 Tolan AM, Kaji AH, Quach C, Hines OJ, de Virgilio C. The electronic residency application service application can predict accreditation council for graduate medical education competency-based surgical resident performance. *J Surg Educ* 2010;67(06):444–448
- 8 Brothers TE, Wetherholt S. Importance of the faculty interview during the resident application process. *J Surg Educ* 2007;64(06):378–385
- 9 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(04):908–914
- 10 Boyse TD, Patterson SK, Cohan RH, et al. Does medical school performance predict radiology resident performance? *Acad Radiol* 2002;9(04):437–445
- 11 Erlandson EE, Calhoun JG, Barrack FM, et al. Resident selection: applicant selection criteria compared with performance. *Surgery* 1982;92(02):270–275
- 12 Fine PL, Hayward RA. Do the criteria of resident selection committees predict residents' performances? *Acad Med* 1995;70(09):834–838
- 13 Khongphatthanayothin A, Chongsrisawat V, Wananukul S, Sanpavat S. Resident recruitment: what are good predictors for performance during pediatric residency training? *J Med Assoc Thai* 2002;85(Suppl 1):S302–S311
- 14 Dirschl DR, Dahners LE, Adams GL, Crouch JH, Wilson FC. Correlating selection criteria with subsequent performance as residents. *Clin Orthop Relat Res* 2002;(399):265–271

- 15 Egol KA, Collins J, Zuckerman JD. Success in orthopaedic training: resident selection and predictors of quality performance. *J Am Acad Orthop Surg* 2011;19(02):72–80
- 16 Harfmann KL, Zirwas MJ. Can performance in medical school predict performance in residency? A compilation and review of correlative studies. *J Am Acad Dermatol* 2011;65(05):1010–1022.e2
- 17 Spitzer AB, Gage MJ, Looze CA, Walsh M, Zuckerman JD, Egol KA. Factors associated with successful performance in an orthopaedic surgery residency. *J Bone Joint Surg Am* 2009;91(11):2750–2755
- 18 Chole RA, Ogden MA. Predictors of future success in otolaryngology residency applicants. *Arch Otolaryngol Head Neck Surg* 2012;138(08):707–712
- 19 Grewal SG, Yeung LS, Brandes SB. Predictors of success in a urology residency program. *J Surg Educ* 2013;70(01):138–143
- 20 Hayden SR, Hayden M, Gamst A. What characteristics of applicants to emergency medicine residency programs predict future success as an emergency medicine resident? *Acad Emerg Med* 2005;12(03):206–210
- 21 Burish MJ, Fredericks CA, Engstrom JW, Tateo VL, Josephson SA. Predicting success: What medical student measures predict resident performance in neurology? *Clin Neurol Neurosurg* 2015;135:69–72
- 22 Metro DG, Talarico JF, Patel RM, Wetmore AL. The resident application process and its correlation to future performance as a resident. *Anesth Analg* 2005;100(02):502–505
- 23 Chen F, Arora H, Martinelli SM, et al. The predictive value of pre-recruitment achievement on resident performance in anesthesiology. *J Clin Anesth* 2017;39:139–144
- 24 Stohl HE, Hueppchen NA, Bienstock JL. Can medical school performance predict residency performance? Resident selection and predictors of successful performance in obstetrics and gynecology. *J Grad Med Educ* 2010;2(03):322–326
- 25 Alterman DM, Jones TM, Heidel RE, Daley BJ, Goldman MH. The predictive value of general surgery application data for future resident performance. *J Surg Educ* 2011;68(06):513–518
- 26 Bowe SN, Laury AM, Gray ST. Associations between otolaryngology applicant characteristics and future performance in residency or practice: a systematic review. *Otolaryngol Head Neck Surg* 2017;156(06):1011–1017
- 27 Daly KA, Levine SC, Adams GL. Predictors for resident success in otolaryngology. *J Am Coll Surg* 2006;202(04):649–654
- 28 Borowitz SM, Saulsbury FT, Wilson WG. Information collected during the residency match process does not predict clinical performance. *Arch Pediatr Adolesc Med* 2000;154(03):256–260
- 29 Black KP, Abzug JM, Chinchilli VM. Orthopaedic in-training examination scores: a correlation with USMLE results. *J Bone Joint Surg Am* 2006;88(03):671–676
- 30 Carmichael KD, Westmoreland JB, Thomas JA, Patterson RM. Relation of residency selection factors to subsequent orthopaedic in-training examination performance. *South Med J* 2005;98(05):528–532
- 31 Gunderman RB, Jackson VP. Are NBME examination scores useful in selecting radiology residency candidates? *Acad Radiol* 2000;7(08):603–606
- 32 Chen B, Zaebs D, Seel L. A macro to calculate kappa statistics for categorizations by multiple raters. Accessed July 23, 2021 at: <https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/155-30.pdf>
- 33 Hu M, Liu B. Mining and summarizing customer reviews. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD-2004), Seattle, Washington 2004: 168–177
- 34 Nielsen FA. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. Proceedings of the ESWC2011 Workshop on ‘Making Sense of Microposts’: Big things come in small packages 2011: 93–98
- 35 Loughran T, McDonald B. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *J Finance* 2011;66(01):35–65
- 36 Mohammad S, Turney P. Crowdsourcing a word-emotion association lexicon. *Comput Intell* 2013;29(03):436–465
- 37 Lin F, Oh SK, Gordon LK, Pineles SL, Rosenberg JB, Tsui I. Gender-based differences in letters of recommendation written for ophthalmology residency applicants. *BMC Med Educ* 2019;19(01):476
- 38 Grova MM, Jenkins FG, Filippou P, et al. Gender bias in surgical oncology fellowship recommendation letters: gaining progress. *J Surg Educ* 2021;78(03):866–874
- 39 Kobayashi AN, Sterling RS, Tackett SA, Chee BW, Laporte DM, Humbyrd CJ. Are there gender-based differences in language in letters of recommendation to an orthopaedic surgery residency program? *Clin Orthop Relat Res* 2020;478(07):1400–1408
- 40 Madera JM, Hebl MR, Martin RC. Gender and letters of recommendation for academia: agent and communal differences. *J Appl Psychol* 2009;94(06):1591–1599
- 41 Trix F, Psenka C. Exploring the color of glass: letters of recommendation for female and male medical faculty. *Discourse Soc* 2003;14(02):191–220
- 42 Powers A, Gerull KM, Rothman R, Klein SA, Wright RW, Dy CJ. Race- and gender-based differences in descriptions of applicants in the letters of recommendation for orthopaedic surgery residency. *JBJS Open Access* 2020;5(03):e20.00023–e20.00023
- 43 Filippou P, Mahajan S, Deal A, et al. The presence of gender bias in letters of recommendations written for urology residency applicants. *Urology* 2019;134(01):56–61
- 44 Lee AG, Golnik KC, Oetting TA, et al. Re-engineering the resident applicant selection process in ophthalmology: a literature review and recommendations for improvement. *Surv Ophthalmol* 2008;53(02):164–176
- 45 Takayama H, Grinsell R, Brock D, Foy H, Pellegrini C, Horvath K. Is it appropriate to use core clerkship grades in the selection of residents? *Curr Surg* 2006;63(06):391–396