Clinical Orthopaedics
and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

Clinical Research

Are Narrative Letters of Recommendation for Medical Students Interpreted as Intended by Orthopaedic Surgery Residency Programs?

Cameron R. Egan MD[1], Jesse Dashe MD[2], Amira I. Hussein PhD[2], Paul Tornetta III MD[2]

Abstract

Background Narrative letters of recommendation are an important component of the residency application process. However, because narrative letters of recommendation are almost always positive, it is unclear whether those reviewing the letters understand the writer's intended strength of support for a given applicant.

Questions/purposes (1) Is the perception of letter readers for narrative letters of recommendation consistent with the intention of the letter's author? (2) Is there inter-reviewer consistency in selection committee members' perceptions of the narrative letters of recommendation?

Methods Letter writers who wrote two or more narrative letters of recommendation for applicants to one university-based orthopaedic residency program for the 2014 to 2015 application cycle were sent a survey linked to a specific letter of recommendation they authored to assess the intended meaning regarding the strength of an applicant. A total of 247 unstructured letters of recommendation and accompanying surveys were sent to their authors, and 157 surveys were returned and form the basis of this study (response percentage 64%). The seven core members of the admissions committee (of 22 total reviewers) at a university-based residency program were sent a similar survey regarding their perception of the letter. To answer our research question about whether letter readers' perceptions about a candidate were consistent with the letter writer's intention, we used kappa values to determine agreement for survey questions involving discrete variables and Spearman correlation coefficients (SCCs) to determine agreement for survey questions involving continuous variables. To answer our research question regarding inter-reviewer consistency among the seven faculty members, we compared the letter readers' responses to each survey question using intraclass correlation coefficients (ICCs).

Results There was a negligible to moderate correlation between the intended and perceived strength of the letters (SCC 0.26 to 0.57), with only one of seven letter readers scoring in the moderate correlation category. When stratifying the applicants into thirds, there was only slight agreement (kappa 0.07 to 0.19) between the writers and reviewers. There were similarly low kappa values for agreement about how the writers and readers felt regarding the candidate matching into their program (kappa 0.14 to 0.30). The ICC for each question among the seven faculty reviewers ranged from poor to moderate (ICC 0.42 to 0.52).

[1]Harvard Combined Orthopaedic Surgery Residency Program, Massachusetts General Hospital, Boston, MA, USA

[2]Department of Orthopaedic Surgery, Boston Medical Center, Boston, MA, USA

C. R. Egan ✉, Harvard Combined Orthopaedic Surgery Residency Program, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114 USA, Email: Egan.cam@gmail.com

Wolters Kluwer

Conclusion Our results demonstrate that the reader's perception of narrative letters of recommendation did not correlate well with the letter writer's intended meaning and was not consistent between letter readers at a single university-based urban orthopaedic surgery residency program.

Clinical Relevance Given the low correlation between the intended strength of the letter writers and the perceived strength of those letters, we believe that other options such as a slider bar or agreed-upon wording as is used in many dean's letters may be helpful.

Introduction

The selection of applicants to postgraduate medical education programs is a multifactorial process that considers both objective and subjective criteria. Resident selection committees across medical and surgical specialties, including orthopaedic surgery, consider the narrative letter of recommendation as an important subjective component of the residency application [3-6, 9, 10, 13, 17, 18, 21, 31, 33, 36]. The 2018 National Resident Matching Program (NRMP) survey found that 89% of orthopaedic surgery program directors found the narrative letter of recommendation to be an important factor when selecting candidates for an interview [25]. Moreover, a widely cited study by Bernstein et al. [5] surveyed applicants to orthopaedic surgery residency programs and found that they considered the narrative letter of recommendation to be the most important part of their application. Because of the weight placed on the narrative letter of recommendation in the context of the residency application, many studies have attempted to define characteristics that make the narrative letter of recommendation more informative [5, 6, 9, 12, 14, 34]. Although the American Orthopaedic Association introduced a standardized letter of recommendation for orthopaedic surgery residency in 2017 [1], a recent study showed grade inflation limits the utility of the standardized recommendation form [20].

There are inherent biases in narrative letters of recommendation. Applicants are likely to choose a letter writer whom they believe will write them a strongly supportive letter, and letter writers do not benefit from potentially portraying their program poorly by writing an unflattering recommendation. The result is that narrative letters of recommendation are almost universally positive, making letters on behalf of more or less competitive applicants harder to distinguish from one another [5-7, 9, 12, 14, 27, 30]. Despite the potential for confidential communication regarding an applicant's strengths and weaknesses, many narrative letters of recommendation miss some or all key factors regarding an applicant's work ethic, interpersonal skills, and teamwork [14, 16, 24, 27]. In addition to the variability in the factors presented in the narrative letter of recommendation, there is also inconsistency in the interpretation of these letters. In 2000, Dirschl and Adams [12] evaluated narrative letter of recommendation inter-reviewer reliability among six orthopaedic faculty by categorizing the letters for 58 applicants as "outstanding," "average," or "poor," and reported kappa scores ranging from 0.17 to 0.28, concluding there was minimal to no agreement among readers. Further studies have established that variability exists in both narrative letter of recommendation presentation and narrative letter of recommendation perception regarding the strength of an applicant across specialties [5, 11, 14, 15, 27]. However, none of the previous studies evaluated the letter writer's intended strength of the letter of recommendation. Therefore, despite the variability in letter reader perception, it is unknown whether most letter readers understand the letter writer's intention. Nonetheless, the variability in narrative letter of recommendation presentation and perception can leave readers feeling as if they are cracking a code to grasp the letter writer's intention. The importance of the narrative letter of recommendation in orthopaedic surgery residency selection underscores the need to better understand the value of the narrative letter of recommendation. However, we found no assessment of whether the intended strength of the recommendation is accurately interpreted by those reading the letters.

We therefore asked: (1) Is the perception of letter readers for narrative letters of recommendation consistent with the intention of the letter's author? (2) Is there inter-reviewer consistency in selection committee members' perceptions of the narrative letters of recommendation?

Materials and Methods

Primary and Secondary Study Outcomes

Our primary study goal was to evaluate whether the perceived strength of a narrative letter of recommendation was consistent with the intended strength by its author. To achieve this, letter writers who wrote two or more narrative letters of recommendation for applicants to a single university-based orthopaedic residency program for one application cycle were sent a survey of six questions linked to a specific letter they authored to assess the intended meaning of the letter (Fig. 1) (Qualtrics Software). In all, 247 narrative letters of recommendation, each with an associated survey, were returned to 99 authors. Of the surveys sent, 64% (157 of 247) of those with completed data for at least one letter of recommendation were returned by 65% (64 of 99) of letter writers.

To evaluate for nonresponse bias, we compared early versus late respondents with the assumption that late respondents were most similar to nonrespondents and found

Q1. Based on your letter of recommendation, please mark how this student ranks:

Rank

Q2. Compared to other candidates you recommended this year, this candidate is in the:
(ie. If you recommended three candidates then one should be represented in each 1/3 category below)

	Rank	# Recommended in this category during academic year
Best this year	☐	
Top 1/3	☐	
Middle 1/3	☐	
Lower 1/3	☐	

Q3. Please mark how well you know this applicant:

Minimal direct experiences with applicant   Worked directly with applicant over extended period of time

Rank

Q4. Please mark how you would feel if this applicant matched into your program:

Slight Disappointment   Not Unhappy   Satisfied   Very Pleased   Ecstatic

Rank

Q5. If you were the one to make the decision, would you want this applicant in your program?

○ Yes
○ No

Q6. Did this applicant match into your program?

○ Yes
○ No

Fig. 1 Survey sent to letter writers along with the letter of recommendation they wrote. A color image accompanies the online version of this article.

no difference in the score variation for the survey questions [35]. We used survey response variation as previous evidence suggests straight-line answering, when survey respondents give the same answer for every question, is an indicator for survey fatigue [28]. We also compared our response rate to similar data gathering surveys of stakeholders within orthopaedic surgery residency programs. For example, the annual post-match survey of orthopaedic

surgery program directors conducted by the NMRP had a response rate of 34.5% in 2018 and 25.9% in 2020 [25, 26].

Letter writers who wrote two or more narrative letters of recommendation were chosen because the writers would be able to compare those letters directly. Further, by using this methodology, we did not select for the quality of the applicant in other areas so that we could evaluate an academically diverse group. The intended strength of the letter was assessed using a slider bar and by asking which third (top, middle, or bottom) of the applicant pool the candidate would fall into. Additionally, the letter writers were asked to evaluate their familiarity with the applicant and most importantly, how happy he or she would be to have the candidate match into their program (also a slider bar). A slider does not force a survey taker into a discrete choice, but rather allows him or her to mark the preference on a continuous scale.

To evaluate whether the author's intended strength was consistent with the reader's perceived strength of the narrative letter of recommendation, seven core admissions committee faculty who had already had access to the letters during the application process were sent a similar survey of five questions (Fig. 2). The seven letter readers were the core faculty on the admissions committee from a medium-sized, urban, university-based program. The cohort consisted of both junior- and senior-level faculty with 2 to 22 years of experience reading and writing letters of recommendation. They were diverse in age (40s to 60s), cultural background, and had experience practicing orthopaedics in a variety of settings including in the United States Armed Forces, private practice, and academic medicine. The survey evaluated the seven faculty members' perceptions of the letter with respect to its strength of recommendation into which third (top, middle, or bottom) the applicant would be assigned, the familiarity of the writer with the applicant, and how satisfied the letter writer would be to have the candidate match into his or her own program. Our secondary goal was to evaluate whether the perceived strength of a narrative letter of recommendation was consistent among letter readers. To achieve these, we evaluated the consistency of responses of the seven faculty members to each survey question for a particular narrative letter of recommendation. The seven faculty members returned a survey for each of the 157 letters they evaluated (response rate 100%). All letters reviewed by the faculty were blinded to the name of the candidate. No candidate who matched into or were students at the university program were included, and the statistics were performed so they were blinded to the letter writers and candidates.

Ethical Approval

Ethical approval for this study was waived by the Boston University Medical Center Institutional Review Board.

Statistical Analysis and Data Treatment

The letter writer intention and letter reader perception surveys contained both continuous and discrete variables. Therefore, letter writer and letter reader answers to survey questions using the slider bar (questions 1, 3, and 4) with continuous variables were compared using Spearman correlation coefficients (SCCs). To compare the letter writer versus letter reader responses to questions with discrete variables (questions 2 and 5), which placed candidates into categorical groupings, a kappa value was determined (JMP Software version 13). The inter-reader agreement in perception for each letter among the seven faculty members was assessed for each question by determining the intraclass correlation coefficient (ICC). The significance level was set to $p < 0.05$. We used the SCCs and kappa values to assess the extent of agreement.

Cutoffs used for SCC values of agreement were 0.0 to 0.3 = negligible; 0.3 to 0.5 = low positive; 0.5 to 0.7 = moderate positive; 0.7 to 0.9 = high positive; and > 0.9 = very high positive [19]. Kappa value determinations were 0.01 to 0.20 = slight; 0.21 to 0.40 = fair; 0.41 to 0.60 = moderate; 0.61 to 0.80 = substantial; and 0.81 to 0.99 = perfect [22]. The ICC of reliability was < 0.5 = poor; 0.5 to 0.75 = moderate; 0.75 to 0.9 = good; and > 0.9 = excellent [29]. Not all of the letter reviewers answered all survey questions; the analysis was completed using all available data. The overall response proportion was 88% (range 57% for Question 2 to 100% for Question 5). Applicants in the "best" category of the letter writer survey were grouped as the top third to allow us to implement categorical variables into the analysis. The question on the letter writer survey regarding whether the applicant had matched into the letter writer's program (Question 6) was not used in our analysis.

Results

Is the Perception of Letter Readers Consistent with the Intention of Letter Writers?

We found a negligible to moderate positive correlation between the letter writer's overall rank of the candidate and the letter reader's perceived rank of the candidate (survey question 1) (SCC 0.26 to 0.57; all p values < 0.01), with only one of seven letter readers scoring in the moderate positive correlation category. When stratifying the applicants into thirds (survey question 2), there was slight agreement (kappa 0.07 to 0.19; all p values < 0.01) between the intended meaning of the writers and the perception of the readers. Readers' perceptions on how well the writer knew the applicant (survey question 3) had a negligible to moderate positive correlation (SCC

Q1. Based on the letter of recommendation, please mark how this student ranks:

Rank

●——————————————————————————————————————

Q2. Compared to other candidates you have reviewed, this candidate is in the: (ie. If you have reviewed three candidates then one should be represented in each 1/3 category below)

	Rank	# Reviwed in this category during academic year
Top 1/3	○	☐
Middle 1/3	○	☐
Lower 1/3	○	☐

Q3. Please mark how well the letter writer knows this applicant:

Minimal direct experiences with applicant   Worked directly with applicant over extended period of time

Rank

●——————————————————————————————————————

Q4. Please mark how you would feel if this applicant matched into your program:

Slight Disappointment   Not Unhappy   Satisfied   Very Pleased   Ecstatic

Rank

●——————————————————————————————————————

Q5. If you were the one to make the decision, would you want this applicant in your program?

○ Yes
○ No

Fig. 2 Survey sent to seven core admissions committee members from a medium-sized, urban, university-based orthopaedic surgery residency program.

0.23 to 0.56; all p values < 0.01) with the letter writer's intended meaning. There was similarly low to moderate positive agreement between writers and readers on question 4 regarding how they would feel if a given candidate matched into their program (SCC 0.34 to 0.51; all p values < 0.01). There was slight to fair agreement (kappa 0.14 to 0.30; all p values < 0.01) regarding whether the letter writer would match the candidate if they were the one to make the decision (survey question 5).

There was no consistent relationship between years of experience reading and writing letters; the faculty member

with the most experience reading narrative letters of recommendation only had a consistent moderate positive agreement with the letter writer on how they would rank the applicant (survey question 1) (SCC 0.57; p < 0.001), how well the writer knew the applicant (survey question 3) (SCC 0.56; p < 0.001), and how the letter writer would feel if the applicant matched into their program (survey question 4) (SCC 0.51; p < 0.001).

Is There Inter-Reader Consistency in Readers' Perceptions of the Narrative Letter of Recommendation?

The ICC for each question among the seven faculty readers ranged from poor to moderate (ICC 0.42 to 0.52; all p values < 0.01). Overall, the ICC (0.47; p < 0.01) was in the poor category, with only the question regarding how the writer would rank the applicant (survey question 1) approaching a moderate ICC (0.49; p < 0.01) and survey question 2 regarding which third the letter writer would place the candidate having a low end of moderate for its ICC (0.52; p < 0.01).

Discussion

Orthopaedic surgery residency selection committees and applicants consider the narrative letter of recommendation an important subjective component of the residency application [5, 25]. However, the opportunity for confidential communication regarding intangible components of an applicant (such as work ethic, interpersonal skills, teamwork, and other such traits) often is vague [14, 16, 24, 27], which leaves the reader feeling as if he or she is trying to crack a code to grasp the letter writer's intention. One prior study attempted to evaluate inter-reader agreement of narrative letters of recommendation using an overall impression of the. letter without evaluating specific parameters; those authors found poor agreement among letter readers [12]. Our study assessed whether the reader of a narrative letter of recommendation has accurately perceived the letter writer's intended meaning and whether there was consistency within the letter readers' perceptions of the narrative letter of recommendation using slider bars and Likert scales. Our results demonstrate that the reader's perception of narrative letters of recommendation did not correlate well with the letter writer's intended meaning. To improve this process, we suggest changes at the institution level, such as program director instructions regarding letters of recommendation and providing a system for letter writers to receive feedback. We also suggest the implementation of national letter of recommendation standards such as agreed-upon wording as is found in many dean's letters

and using slider bars, which allows the letter writers to score candidates on a continuous scale.

Limitations

There are several limitations that may have influenced our results. The present study included seven readers from an admissions committee from one university-based urban program. The composition of this group included all of the full-time faculty members on the admissions committee. The group was diverse in its orthopaedic practice background, years of practice, age (40s to 60s), cultural heritage, and years of experience reading and writing letters (2 to 22 years). However, all seven were men. Although not uncommon in orthopaedics [8], the fact there were no women in this group may limit the generalizability of our study. Furthermore, the letter readers were from a single residency program, which, like all programs, has a distinct culture with specific ideas regarding what constitutes a highly desirable candidate for their program. For example, letter readers from a residency program with a research focus may rank a candidate higher when a letter highlights an applicant's research acumen. Therefore, letter writers may have indicated that they had intended to write a strong letter for a candidate, but our letter readers perceived those letters in the context of what they value in a resident at a single program. However, a distinct culture is true of every program. Furthermore, our letter reader survey also included a question regarding how well the author knew the applicant, which we assumed would be free from the previously mentioned cultural influence, and the letter readers remained in the negligible to moderate positive correlation category (SCC 0.21 to 0.56; all p values < 0.01). Therefore, we believe the culture of our letter readers' program had minimal influence on our results.

Our study may also be limited by nonresponse bias as 35% (35 of 99) of the letter writers who received surveys did not respond. Although we did not perform a follow-up survey with nonrespondents to discover why they did not participate, we evaluated nonresponse bias by comparing the variability in survey data from early versus late respondents and found no difference. We also compared our response rate to similar data gathering surveys of stakeholders within orthopaedic surgery residency programs. For example, the NRMP annual post-match survey of orthopaedic surgery program directors had a response rate of 34.5% in 2018 and 25.9% in 2020 [25, 26]. Given the lack of variability in our early versus late respondents and our response rate compared with previous surveys with a similar population, we do not believe that our respondents were substantially different than those who did not

respond. Therefore, we do not believe nonresponse bias influenced the quality of the letter writer data obtained.

It is also important to recognize how recall bias may have influenced our results. The narrative letters of recommendation used for the present study were written for candidates applying in the 2014 to 2015 application cycle and were returned to their authors in 2017. Given the length of time between letter writing and survey completion, recall bias may have impacted the letter writer's recollection of a candidate's strengths and weaknesses. To minimize this potential for bias, we returned the specific letter of recommendation the letter writer had authored along with the corresponding survey. Therefore, the letter writer was able to reference their own letter while completing the survey.

Is the Perception of Letter Readers Consistent with the Intention of Letter Writers?

The present study demonstrates that the narrative letter of recommendation did not consistently convey the author's intended meaning to the letter reader. These findings may be explained by the consistent inflation of positive statements within the letters. The letter writer has little to gain by portraying a candidate as average, even when this may be the case. Furthermore, standout terms may be hard to distinguish unless their meaning is agreed-upon and consistent. To improve the letter reader's understanding of a letter writer's intended meaning, we believe further work needs to focus on administrative body guidelines for the standardized letter of recommendation as well as continuing education for letter writers and letter readers. Specifically, we suggest the Council of Orthopaedic Residency Directors (CORD) implement guidelines regarding key adjectives with agreed-upon definitions similar to the Association of American Medical Colleges (AAMC) guideline for the dean's letter of evaluation [2], which uses a single adjective to describe a candidate's standing, namely Outstanding (top 10%), Excellent (top 30%), Very Good (top 50%), Good (top 70%), and Fair (top 100%). To make this adjective meaningful, the AAMC guidelines state that it should only be used when an adequate comparison group is provided, thereby giving the rating system context [2]. We suggest CORD implement similar guidelines that ask the letter writer to provide context in comparing the student he or she is recommending to other medical students applying to orthopaedics from their institution. We also believe continuing education sessions during annual meetings of the American Academy of Orthopaedic Surgeons and CORD could provide information and instruction focusing on clear and effective communication. At the institution level, we suggest program directors provide their faculty with clear guidelines regarding how their medical students compare with one another, and that they actively limit inflation. We also believe that letter writers rarely receive feedback regarding their letters of recommendation. Previous authors have suggested the implementation of a feedback system for letter writers [23], and we believe providing feedback to letter writers both through their institution and directly from letter readers will improve the quality of letters of recommendation overall. Finally, using a slider bar to gauge how pleased a writer would be if the candidate matched into his or her program that includes mostly positive terms used may be helpful. As an example, a scale that went from "slightly disappointed" to "happy" and ending in "thrilled", with other terms in between, might make it more palatable for a writer to place someone at or near the "bottom" of the scale.

Is There Inter-Reader Consistency in Readers' Perceptions of the Narrative Letter of Recommendation?

The inter-reader agreement among the seven faculty members' perceptions was, at best, moderate, which is not high enough to be useful. We anticipated a higher inter-reader consistency for our study given that we asked our readers to comment on the specific components of a letter. Although we asked more specific questions than did Dirschl and Adams [12], who showed slight interobserver reliability among their faculty when categorizing narrative letters of recommendation into poor, average, or outstanding, we believe our letter readers found the almost universally positive letters difficult to distinguish. Previous investigations of standardized letters of recommendation in other specialties showed that the standardized letter of recommendation has better inter-reader reliability and takes less time to interpret than narrative letters of recommendation [15, 36]. To ensure candidates are assessed accurately, CORD has adopted the American Orthopaedic Association's standardized letter of recommendation [1]. However, similar problems with almost universally positive scores for each candidate may negate the ease of analyzing objective data provided in the standardized format, with a recent study showing 92% of candidates were ranked in the top two categories [20, 32]. Furthermore, a recent retrospective study showed that 24% of standardized letter of recommendation authors provided additional supplemental narrative statements to their standardized evaluation [32]. Therefore, even in the standardized letter of recommendation, effective communication in the narrative format remains important. To improve the communication of accurate information within the standardized letter of recommendation, we believe changing the scale would improve the

willingness of authors to differentiate students. Therefore, we propose adding a question to the standardized letter of recommendation as described above with a slider bar regarding how happy the author would be to have the candidate in his or her program, with "happy" in the middle of the bar, as opposed to "middle third." Further studies should focus on which questions on the standardized letter of recommendation are the most important to the letter readers in their assessment of a candidate.

Conclusion

Our results demonstrate that the reader's perception of narrative letters of recommendation did not correlate well with the letter writer's intended meaning. Given the low correlation between the intended strength of the letter writers and the perceived strength of those letters, we believe that other options such as agreed-upon wording, like that found in many dean's letters, and slider bars, which allows the letter writer to assess a candidate on a continuous scale, may be useful adjuncts.

References

1. American Orthopaedic Association. Standardized letter of recommendation. Available at: https://2wq9z720przc30yf331grssd-wpengine.netdna-ssl.com/wp-content/uploads/2020/12/AOA-CORD-SLOR-2017-V8161.pdf. Accessed November 10, 2018.
2. Association of American Medical Colleges. MSPE Task Force recommendations. Available at: https://www.aamc.org/professional-development/affinity-groups/gsa/medical-student-performance-evaluation. Accessed November 16, 2020.
3. Bajaj G, Carmichael KD. What attributes are necessary to be selected for an orthopaedic surgery residency position: perceptions of faculty and residents. *South Med J*. 2004;97:1179-1185.
4. Belmont PJ Jr, Hoffmann JD, Tokish JM, et al. Overview of the military orthopaedic surgery residency application and selection process. *Mil Med*. 2013;178:1016-1023.
5. Bernstein AD, Jazrawi LM, Elbeshbeshy B, Della Valle CJ, Zuckerman JD. An analysis of orthopaedic residency selection criteria. *Bull Hosp Jt Dis*. 2002;61:49-57.
6. Bernstein AD, Jazrawi LM, Elbeshbeshy B, Della Valle CJ, Zuckerman JD. Orthopaedic resident-selection criteria. *J Bone Joint Surg Am*. 2002;84:2090-2096.
7. Caldroney RD. Letters of recommendation. *J Med Educ*. 1983;58:757-758.
8. Chambers CC, Ihnow SB, Monroe EJ, Suleiman LI. Women in orthopaedic surgery: population trends in trainees and practicing surgeons. *J Bone Joint Surg Am*. 2018;100:e116.
9. Clark R, Evans EB, Ivey FM, Calhoun JH, Hokanson JA. Characteristics of successful and unsuccessful applicants to orthopedic residency training programs. *Clin Orthop Relat Res*. 1989;(241):257-264.
10. DeLisa JA, Jain SS, Campagnolo DI. Factors used by physical medicine and rehabilitation residency training directors to select their residents. *Am J Phys Med Rehabil*. 1994;73:152-156.
11. DeZee KJ, Magee CD, Rickards G, et al. What aspects of letters of recommendation predict performance in medical school? Findings from one institution. *Acad Med*. 2014;89:1408-1415.
12. Dirschl DR, Adams GL. Reliability in evaluating letters of recommendation. *Acad Med*. 2000;75:1029.
13. Egol KA, Collins J, Zuckerman JD. Success in orthopaedic training: resident selection and predictors of quality performance. *J Am Acad Orthop Surg*. 2011;19:72-80.
14. Fortune JB. The content and value of letters of recommendation in the resident candidate evaluative process. *Curr Surg*. 2002;59:79-83.
15. Girzadas DV Jr, Harwood RC, Dearie J, Garrett S. A comparison of standardized and narrative letters of recommendation. *Acad Emerg Med*. 1998;5:1101-1104.
16. Girzadas DV Jr, Harwood RC, Delis SN, et al. Emergency medicine standardized letter of recommendation: predictors of guaranteed match. *Acad Emerg Med*. 2001;8:648-653.
17. Gordon MJ, Lincoln JA. Selecting a few residents from many applicants: a new way to be fair and efficient. *J Med Educ*. 1976;51:454-460.
18. Greenburg AG, Doyle J, McClure DK. Letters of recommendation for surgical residencies: what they say and what they mean. *J Surg Res*. 1994;56:192-198.
19. Hinkle DE, Wiersma W, Jurs SG. *Applied Statistics for the Behavioral Sciences*. Houghton Mifflin; 2003.
20. Kang HP, Robertson DM, Levine WN, Lieberman JR. Evaluating the standardized letter of recommendation form in applicants to orthopaedic surgery residency. *J Am Acad Orthop Surg*. 2020;28:814-822.
21. LaGrasso JR, Kennedy DA, Hoehn JG, Ashruf S, Przybyla AM. Selection criteria for the integrated model of plastic surgery residency. *Plast Reconstr Surg*. 2008;121:121e-125e.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
23. Leopold SS. Editor's spotlight/take 5: Are there gender-based differences in language in letters of recommendation to an orthopaedic surgery residency program? *Clin Orthop and Relat Res*. 2020;478:1396-1399.
24. Marwan Y, Waly F, Algarni N, et al. The role of letters of recommendation in the selection process of surgical residents in Canada: a national survey of program directors. *J Surg Educ*. 2017;74:762-767.
25. National Resident Matching Program DRaRC. Results of the 2018 NRMP program director survey. Available at: https://www.nrmp.org/wp-content/uploads/2018/07/NRMP-2018-Program-Director-Survey-for-WWW.pdf. Accessed December 1, 2018.
26. National Resident Matching Program DRaRC. Results of the 2020 NRMP program director survey. 2020. Available at: https://mk0nrmp3oyqui6wqfm.kinstacdn.com/wp-content/uploads/2020/08/2020-PD-Survey.pdf. Accessed November 20, 2020.
27. O'Halloran CM, Altmaier EM, Smith WL, Franken EA Jr. Evaluation of resident applicants by letters of recommendation: a comparison of traditional and behavior-based formats. *Invest Radiol*. 1993;28:274-277.
28. O'Reilly-Shah VN. Factors influencing healthcare provider respondent fatigue answering a globally administered in-app survey. *PeerJ*. 2017;5:e3785.
29. Portney LG. *Foundations of Clinical Research: Applications to Practice*. Prentice Hall; 2000.
30. Rim Y. How reliable are letters of recommendation. *J High Educ*. 1976;47:437-446.

31. Ross CA, Leichner P. Criteria for selecting residents: a reassessment. *Can J Psychiatry*. 1984;29:681-686.

32. Samade R, Balch Samora J, Scharschmidt TJ, Goyal KS. Use of standardized letters of recommendation for orthopaedic surgery residency applications: a single-institution retrospective review. *J Bone Joint Surg Am*. 2020;102:e14.

33. Sherry E, Mobbs R, Henderson A. Becoming an orthopaedic surgeon: background of trainees and their opinions of selection criteria for orthopaedic training. *Aust N Z J Surg*. 1996;66:473-477.

34. Stohl HE, Hueppchen NA, Bienstock JL. The utility of letters of recommendation in predicting resident success: can the ACGME competencies help? *J Grad Med Educ*. 2011;3:387-390.

35. Voigt LF, Koepsell TD, Daling JR. Characteristics of telephone survey respondents according to willingness to participate. *Am J Epidemiol*. 2003;157:66-73.

36. Wagoner NE, Suriano JR. Program directors' responses to a survey on variables used to select residents in a time of change. *Acad Med*. 1999;74:51-58.

Wolters Kluwer