Journal Articles

2018

Characterization of the 2016-2017 dermatology standardized letter of recommendation

R. F. Wang

M. Zhang

A. Alloo
Zucker School of Medicine at Hofstra/Northwell

T. Stasko

J. E. Miller

See next page for additional authors

Recommended Citation

Wang RF, Zhang M, Alloo A, Stasko T, Miller JE, Kaffenberger JA. Characterization of the 2016-2017 dermatology standardized letter of recommendation. . 2018 Jan 01; 11(3):Article 3379 [p.]. Available from: https://academicworks.medicine.hofstra.edu/articles/3379. Free full text article.

Authors
R. F. Wang, M. Zhang, A. Alloo, T. Stasko, J. E. Miller, and J. A. Kaffenberger

Characterization of the 2016–2017 Dermatology Standardized Letter of Recommendation

by REBECCA F. WANG, BA; MYRON ZHANG, MD; ALLIREZA ALLOO, MD; THOMAS STASKO, MD; JORDAN E. MILLER, BA; and JESSICA A. KAFFENBERGER, MD

Ms. Wang and Drs. Zhang and Kaffenberger are with the Division of Dermatology, Department of Internal Medicine, The Ohio State University Wexner Medical Center in Columbus, Ohio. Dr. Alloo is with the Department of Dermatology at Hofstra North-well School of Medicine in New Hyde Park, New York. Dr. Stasko and Mr. Miller are with the Department of Dermatology at the University of Oklahoma in Oklahoma City, Oklahoma.

ABSTRACT

Objective: We aimed to analyze the reformatted standard letter of recommendation (SLOR) for dermatology residents to examine trends in grading and content based on the positions of the letter writers, their backgrounds, and their relationship with the applicant, as well as to evaluate the SLOR's ability to discriminate applicants. **Design:** This was a retrospective characterization study of dermatology SLORs from the 2016–17 application cycle. **Setting:** We examined SLORs received by The Ohio State University, the University of Oklahoma, and Hofstra University Northwell Health dermatology residency programs. **Participants:** We included dermatology residency applicants and their letter writers from the 2016–17 application cycle. **Results:** A total of 141 SLORs were analyzed from 115 applicants. SLORs demonstrated grade inflation from letter writers of all backgrounds. Ratings for research potential and inquisitive nature were significantly lower than ratings for other categories. Letter writers with limited clinical and research contact graded applicants significantly lower than did writers who had more extensive contact. Word boxes were underutilized. **Conclusion:** The dermatology SLOR is useful in differentiating applicants, and ratings correlate with the relationships that letter writers have with their applicants. Residency programs should be aware of these findings when evaluating letters of recommendation for applicants.

KEYWORDS: Letter of recommendation, standardized, residency, application, applicant

Traditionally, the narrative letter of recommendation (NLOR) has been used by dermatology residency programs to gather useful information regarding the personality and character strengths of the applicants. However, due to the free-form nature of NLORs, the letters can be excessively flattering,[1–3] lack clarity,[3,4] and demonstrate low reliability between interpreting faculty members.[5,6] In an effort to improve the efficiency, validity, and stratification of applicants, a dermatology standardized letter of recommendation (SLOR) was created and first utilized during the 2014–15 application cycle.[7]

SLORs from the 2014–15 application cycle were analyzed, suggesting that they were easier and quicker to interpret, had less exaggeration of applicants' positive traits, and demonstrated higher interrater and intrarater reliability compared to NLORs.[7] However, numerous letter writers felt that the 2014–15 SLOR had several weaknesses, so an American Academy of Dermatology work group reformatted the SLOR for the 2016–17 application cycle. The reformatted SLOR has greater space for descriptive feedback as well as modified grading categories. The primary aim of this study was to analyze the reformatted SLOR to examine trends in grading and content based on letter writer position, background, and relationship with the applicant, as well as to evaluate the SLOR's ability to discriminate applicants.

METHODS

Study design. This was a retrospective, multi-institutional study of SLORs received by The Ohio State University, the University of Oklahoma, and Hofstra University Northwell Health dermatology residency programs during the 2016–17 application cycle. Institutional review board approval was obtained from each participating institution (2016B0466). Only reformatted SLORs were analyzed; any SLORs of the older format were excluded from this study. Duplicate letters were removed, and all letters were de-identified. All responses from each SLOR were compiled, and word counts were obtained for free text boxes.

Standardized letter of recommendation. Letter writers were asked to select their present position and identify the number of years they had been in that position. If multiple positions were indicated, only the highest academic rank was considered, with the highest rank being department chair, followed by program director, assistant program director, dermatology faculty, non-dermatology faculty, research faculty, and private practice physician. Letter writers also answered a series of background questions regarding their contact with medical students

in general and with the applicant specifically. In the final section of the SLOR, letter writers assessed the applicant in comparison to the overall dermatology applicant pool with respect to the following categories: work ethic, self-initiative, dependability, ability to work as part of a team, communication skills, research potential, and inquisitive nature. For each category, letter writers were asked to select a rating of "not enough exposure," "below average," "average," "excellent (top 33%)," "outstanding (top 15%)," or "exceptional (top 5%)." For each category, a comment box was provided with a word limit of 50 words. Finally, letter writers were asked to indicate the applicant's greatest strength and include any additional comments in a word box with a 200-word limit.

Statistical analysis. Analysis was performed using MATLAB (version R2015b; MathWorks, Natick, MA, USA). The frequency of each assessment rating was calculated across all question categories and stratified by respondent characteristics. A few instances of skipped responses were omitted from the analysis. To evaluate differences in assessment ratings across respondent categories (e.g., position, contact with applicants), each rating was converted to a numerical rank (0=below average, 1=average, 2=excellent, and so on). A mean rank across question categories was calculated for each respondent, and the Kruskal-Wallis (KW) test was performed to evaluate any differences in the distributions of ratings between respondent groups. If there was a statistically significant difference, pairwise *post-hoc* analysis was performed using the Tukey-Kramer method to correct for multiple comparison. *P* values less than 0.05 were considered to be statistically significant.

RESULTS

A total of 141 SLORs were analyzed from 115 unique applicants across three institutions. Ninety out of 472 applicants used one or more SLORs when applying to The Ohio State University (19%), while 77 out of 508 applicants used SLORs for Hofstra University Northwell Health (15%) and and 73 out of 442 applicants used SLORs for the University of Oklahoma (17%). The most common authors were program directors, followed by dermatology faculty, department chairs, and assistant program directors. Table 1 provides a breakdown of the average length of time letter writers have held their position and known the applicant. No letters were written by non-

TABLE 1. Letter writer background		
POSITION	**AVERAGE YEARS AT POSITION (SD)**	**AVERAGE NUMBER OF MONTHS OF KNOWING APPLICANT (SD)**
Department chair	8.0 (9.1)	13.1 (13.2)
Program director	7.8 (5.7)	14.6 (15.2)
Assistant program director	2.6 (0.9)	4.5 (6.7)
Dermatology faculty	9.2 (10.8)	11.3 (12.9)
TOTAL	**7.8 (8.4)**	**12.2 (13.5)**
SD: standard deviation		

dermatology faculty, research faculty, or private practice physicians. Three letters were written by dermatopathologists, all of whom indicated they were dermatology faculty members.

The percentage of applicants receiving each rating based on letter writers' present position, number of medical students worked with in the past year, how often they work with medical students per week, and their degree of contact with the applicant are shown in Table 2 and Figure 1. The KW test did not detect any significant differences in rating distributions by position (*P*=0.82), number of students (*P*=0.80), or days per week working with students (*P*=0.63), but did report a difference for degree of contact with the applicant (*P*=0.0019). Tukey-Kramer showed that letter writers with limited clinical contact gave lower ratings compared to those with extended research contact (*P*=0.018) and those who observed the applicant writing an article (*P*=0.049). Ratings from letter writers with limited research contact were also lower than those from letter writers with extended research contact (*P*=0.046).

The percentage of 50-word limit boxes used and the average word count per box are also shown in Table 2. None of the 50-word limit boxes were used in 52 of the SLORs (37%). The recommended word count for the 50-word limit boxes was exceeded two percent of the time (20 out of 987 word boxes). The recommended word count for the 200-word limit boxes was exceeded 26 percent of the time (36 out of 141 letters), with an average word count of 158 words (standard deviation [SD}: 75). When a grade of "exceptional" was given, the word box was used 70 percent of the time as compared to with "outstanding" (35% of the time), "excellent" (35% of the time), "average" (34% of the time), and "not enough exposure" (14% of the time).

Rating frequencies for each question category are shown in Table 3 and Figure 1. A "below

average" rating was given by one letter writer for research potential and was not selected at all for the six remaining categories. From the Tukey-Kramer test, ratings for research potential were significantly lower than those for all other categories (*P*<0.032) except inquisitive nature. Ratings for inquisitive nature were lower than those for both work ethic (*P*=0.037) and self-initiative (*P*=0.021).

DISCUSSION

This study analyzed the dermatology SLOR from the 2016–17 application cycle to examine the SLOR's ability to differentiate applicants, as well as how backgrounds of the letter writers and their relationships with the applicants influence grading. Analysis of SLORs demonstrated grade inflation from letter writers of all backgrounds, with only one letter writer giving a "below average" grade and a grade of "average" used just 9.3 percent of the time (Table 2). An "exceptional" grade, which as specified by the SLOR should be reserved for the top five percent of applicants, was given 25.4 percent of the time (Table 2). Given past studies that demonstrated hesitation by dermatologists in academia to address negative qualities of applicants,[8] applicants' selection of letter writers who will write them favorable letters,[9] and NLORs that contain only positive feedback about the applicant,[10] grade inflation with the SLOR might be difficult to eradicate. Additionally, it is possible that grade inflation is more marked with the initial use of the SLOR, as letter writers do not want to hurt the chances of their students at matching by utilizing a new letter format.

Despite the presence of grade inflation, the SLOR demonstrated a range of responses suggestive of a differentiation among candidates. The greatest range of responses as well as the most "average" ratings were given for research potential and inquisitive nature (Table 3). A

TABLE 2. Rating frequencies and word box use based on letter writer background

FACTORS	LETTER WRITERS n (%)	RATING PERCENTAGES					% WORD BOXES USED *	AVERAGE WORD COUNT M (SD) *
		NOT ENOUGH EXPOSURE	AVERAGE	EXCELLENT	OUTSTANDING	EXCEPTIONAL		
HIGHEST ACADEMIC RANK								
Department chair	37 (26.2)	2.7	5.4	27.9	41.1	22.5	36.7	31.1 (26.1)
Program director	49 (34.8)	2.3	10.2	30.0	34.7	22.7	39.1	19.6 (12.7)
Assistant program director	13 (9.2)	1.1	5.5	22.0	47.3	24.2	42.9	29.6 (11.1)
Dermatology faculty	42 (29.8)	2.1	13.0	23.3	30.1	31.5	54.4	26.4 (18.0)
MEDICAL STUDENTS WORKED WITH IN THE PAST YEAR								
Less than one per week	4 (2.9)	0.0	28.6	14.3	17.9	39.3	42.9	24.7 (10.2)
One to three per week	35 (25.2)	2.4	6.5	23.3	42.0	25.3	34.3	26.8 (29.0)
More than three per week	100 (71.9)	2.3	9.8	28.7	35.3	24.0	47.0	25.3 (15.5)
CONTACT WITH THE APPLICANT								
Limited clinical	32 (22.7)	2.2	14.7	30.8	34.8	17.0	30.8	22.5 (13.4)
Limited research	12 (8.5)	3.6	14.3	27.4	51.2	3.6	36.9	17.8 (8.1)
Extended clinical	90 (63.8)	2.1	7.5	26.6	37.8	26.0	49.8	26.9 (20.3)
Extended research	35 (24.8)	1.2	5.3	17.3	32.1	44.0	51.0	29.8 (19.3)
Writing an article	50 (35.5)	1.4	6.1	18.2	39.2	35.2	49.4	29.4 (24.8)
Program director	35 (24.8)	2.5	8.6	25.8	40.2	23.0	48.6	20.0 (13.2)
Advisor	41 (29.1)	2.5	6.0	19.6	37.5	34.4	42.2	26.6 (19.8)
TOTAL		**2.2**	**9.3**	**26.7**	**36.2**	**25.4**	**43.4**	**25.6 (18.7)**

n: number; M: mean; SD: standard deviation
*Word boxes used and average word count are based on 50-word limit boxes only. Average word counts include only word boxes with word count > 0.
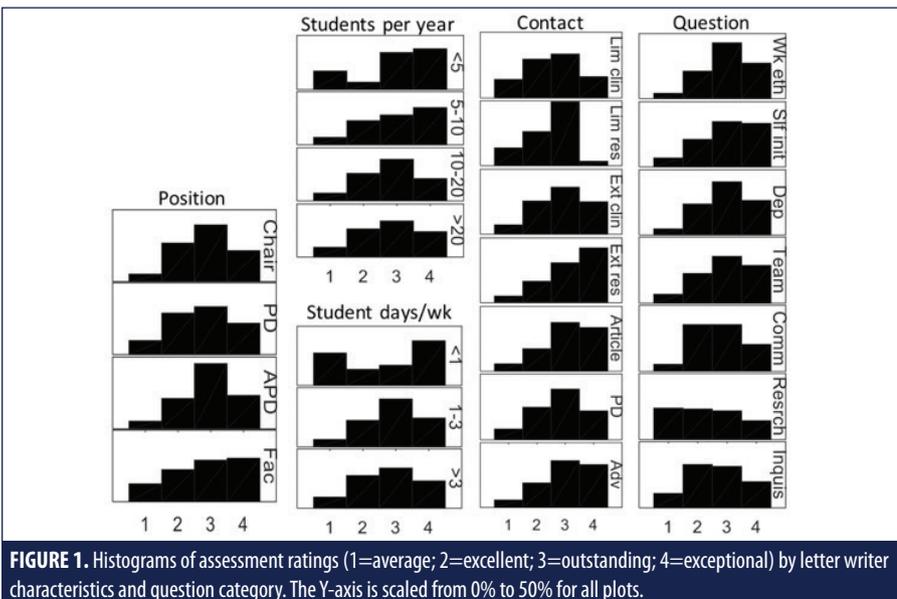

FIGURE 1. Histograms of assessment ratings (1=average; 2=excellent; 3=outstanding; 4=exceptional) by letter writer characteristics and question category. The Y-axis is scaled from 0% to 50% for all plots.

previous analysis of otolaryngology SLORs found that research did not correlate with successful matches, whereas interpersonal and communication skills, initiative and drive, and match potential had a significant association with matching.[11] Thus, while it might simply be easier to stratify applicants based on research potential and inquisitive nature, it is possible that writers believe that lower ratings in these two categories are less likely to negatively impact a candidate's application.

We also examined the SLOR's ability to differentiate top applicants from the rest of the applicant pool. "Exceptional" was given as a rating 25.4 percent of the time; however, it was not the most frequent grade for any of the seven question categories. Furthermore, the word box was used for an "exceptional" grade 70 percent of

TABLE 3. Rating frequencies for each question category of the standardized letter of recommendation

APPLICANT QUALITY	RATING				
	NOT ENOUGH EXPOSURE, n (%)	AVERAGE, n (%)	EXCELLENT, n (%)	OUTSTANDING, n (%)	EXCEPTIONAL, n (%)
Work ethic	0 (0.0)	6 (4.3)	31 (22.1)	63 (45.0)	40 (28.6)
Self-initiative	0 (0.0)	10 (7.1)	31 (22.0)	51 (36.2)	49 (34.8)
Dependability	0 (0.0)	7 (5.0)	35 (24.8)	60 (42.6)	39 (27.7)
Teamwork	4 (2.8)	10 (7.1)	33 (23.4)	52 (36.9)	42 (29.8)
Communication	0 (0.0)	8 (5.7)	51 (36.2)	52 (36.9)	30 (21.3)
Research potential	17 (12.1)	35 (25.0)	34 (24.3)	32 (22.9)	21 (15.0)
Inquisitive nature	1 (0.7)	16 (11.4)	48 (34.3)	46 (32.9)	29 (20.7)

n: number

the time, falling short of the SLOR requirement, "If this candidate is below average or exceptional in any of the areas, please include further details following each category." While our study was not designed to determine whether the "exceptional" category appropriately corresponds to the top quartile of students, we believe that greater use of word boxes, particularly in this category, could help readers differentiate students in this top tier.

Upon analysis of the degree of contact the letter writers had with their applicants, we found that letter writers with more limited clinical or research contact graded applicants significantly lower than did those writers who had more extensive contact. Beskind et al[12] similarly demonstrated that emergency medicine SLOR writers who knew the students for more than one year assigned higher scores as compared with those who knew the students for less than one year.[12] As one might expect, it appears that letter writers who have worked more closely with their applicants tend to give higher ratings.

Finally, analysis of the narrative text demonstrated underuse of the word boxes. Nearly 40 percent of the writers chose not to give any description after each question category, and only 70 percent of writers utilized the required word box after selecting "exceptional." Future studies are needed to examine whether increased use of the word boxes could help to control grade inflation.

Limitations. Our study examined SLORs from three institutions. Although the institutions selected were geographically diverse in an effort to capture SLORs from applicants across the country, it is possible that we did not analyze all SLORs used in the 2016–2017 application cycle. Furthermore, because we did not analyze data from applicants who used only NLORs, the trends noted in this study might not be representative of all dermatology applicants and letter writers. Finally, we did not examine other components of the residency application such as Step 1 score, Step 2 score, and American Osteopathic Association membership in relation to SLOR grading. As this was solely a descriptive study, we are unable to draw conclusions about the reliability of the SLOR.

CONCLUSIONS

Despite grade inflation, the SLOR has utility in differentiating applicants. The categories of research potential and inquisitive nature demonstrated the largest range of responses. SLORs from letter writers who had less contact with the applicants were less inflated. Finally, the narrative sections were underutilized in the SLOR format. Residency programs should be aware of these findings as they evaluate letters of recommendation from applicants.

REFERENCES

1. Fortune JB. The content and value of letters of recommendation in the resident candidate evaluative process. *Curr Surg.* 59(1):79–83.
2. Morgenstern BZ, Zalneraitis E, Slavin S. Improving the letter of recommendation for pediatric residency applicants: an idea whose time has come?. *J Pediatr.* 2003;143(2):143–144.
3. Keim SM, Rein JA, Chisholm C, et al. A standardized letter of recommendation for residency application. *Acad Emerg Med.* 1999;6(11):1141–1146.
4. DeZee KJ, Thomas MR, Mintz M, Durning SJ. Letters of recommendation: rating, writing, and reading by clerkship directors of internal medicine. *Teach Learn Med.* 21(2):153–158.
5. Dirschl DR, Adams GL. Reliability in evaluating letters of recommendation. *Acad Med.* 2000;75(10):1029.
6. Greenburg AG, Doyle J, McClure DK. Letters of recommendation for surgical residencies: what they say and what they mean. *J Surg Res.* 1994;56(2):192–198.
7. Kaffenberger JA, Mosser J, Lee G, et al. A retrospective analysis comparing the new standardized letter of recommendation in dermatology with the classic narrative letter of recommendation. *J Clin Aesthet Dermatol.* 2016;9(9):36–42.
8. Kaffenberger BH, Kaffenberger JA, Zirwas MJ. Academic dermatologists' views on the value of residency letters of recommendation. *J Am Acad Dermatol.* 2014;71(2):395–396.
9. Love JN, Deiorio NM, Ronan-Bentle S, et al. Characterization of the Council of Emergency Medicine Residency Directors' standardized letter of recommendation in 2011-2012. *Acad Emerg Med.* 2013;20(9):926–932.
10. Messner AH, Shimahara E. Letters of recommendation to an otolaryngology/head and neck surgery residency program: their function and the role of gender. *Laryngoscope.* 2008;118(8):1335–1344.
11. Kimple AJ, McClurg SW, Del Signore AG, et al. Standardized letters of recommendation and successful match into otolaryngology. *Laryngoscope.* 2016;126(5):1071–1076.
12. Beskind DL, Hiller KM, Stolz U, et al. Does the experience of the writer affect the evaluative components on the standardized letter of recommendation in emergency medicine?. *J Emerg Med.* 2014;46(4):544–550. JCAD