

Analysis of the Evaluative Components on the Standard Letter of Recommendation (SLOR) in Emergency Medicine

Kristi H. Grall, MD, MHPE
Katherine M. Hiller, MD, MPH
Lisa R. Stoneking, MD

University of Arizona, Department of Emergency Medicine, Tucson, Arizona

Supervising Section Editor: Michael Epter, DO

Submission history: Submitted February 8, 2013; Revision received January 21, 2014; Accepted February 1, 2014

Electronically published May 15, 2014

Full text available through open access at http://escholarship.org/uc/uciem_westjem

DOI: 10.5811/westjem.2014.2.19158

Introduction: The standard letter of recommendation in emergency medicine (SLOR) was developed to standardize the evaluation of applicants, improve inter-rater reliability, and discourage grade inflation. The primary objective of this study was to describe the distribution of categorical variables on the SLOR in order to characterize scoring tendencies of writers.

Methods: We performed a retrospective review of all SLORs written on behalf of applicants to the three Emergency Medicine residency programs in the University of Arizona Health Network (i.e. the University Campus program, the South Campus program and the Emergency Medicine/Pediatrics combined program) in 2012. All “Qualifications for Emergency Medicine” and “Global Assessment” variables were analyzed.

Results: 1457 SLORs were reviewed, representing 26.7% of the total number of Electronic Residency Application Service applicants for the academic year. Letter writers were most likely to use the highest/most desirable category on “Qualifications for EM” variables (50.7%) and to use the second highest category on “Global Assessments” (43.8%). For 4-point scale variables, 91% of all responses were in one of the top two ratings. For 3-point scale variables, 94.6% were in one of the top two ratings. Overall, the lowest/least desirable ratings were used less than 2% of the time.

Conclusions: SLOR letter writers do not use the full spectrum of categories for each variable proportionately. Despite the attempt to discourage grade inflation, nearly all variable responses on the SLOR are in the top two categories. Writers use the lowest categories less than 2% of the time. Program Directors should consider tendencies of SLOR writers when reviewing SLORs of potential applicants to their programs. [West J Emerg Med. 2014;15(4):419–423.]

INTRODUCTION

Background and Importance

Medical student applicants to emergency medicine (EM) residency training programs are required to supply letters of recommendation with their applications through the Electronic Residency Application Service (ERAS), an online service that transmits applications electronically from medical students to residency programs. Applicants are evaluated by residency programs based on various components of their application including United States Medical Licensing Examination (USMLE) scores, the dean’s performance evaluation, clinical

rotation grades, extracurricular experiences, the medical school’s reputation, and letters of recommendation.¹⁻⁴

In 1996, the Council of Residency Directors in Emergency Medicine (CORD) developed a Standard Letter of Recommendation (SLOR) in an attempt to normalize the evaluation of applicants, improve inter-rater reliability of letters of recommendation and to discourage the “upward creep of superlatives.”^{5,6} The SLOR includes student evaluation on the following categorical variables: commitment to EM (CEM), work ethic (WET), ability to develop a treatment plan (DTP), ability to interact with others (IWO),

ability to communicate with patients (CWP), guidance predicted during residency (GUI), prediction of success (PRS), global assessment score (GAS), and likelihood of matching assessment (LOMA). Each variable rates students on a three or four point categorical scale that includes anchors such as “outstanding,” “excellent,” and “good”. Despite widespread use and expectation in EM, the validity of the SLOR has not been well studied, and functional responses to the SLOR categorical variables have not been well characterized.

While the CORD EM has chosen to revise the format of the SLOR to the Standardized Letter of Evaluation (SLOE), most of the categories of the SLOE correspond directly to those of the SLOR. The changes made in revision of the SLOR to the SLOE reflect a greater emphasis on evaluation in addition to recommendation, and a simplification of the form in order to promote standardization across institutions.

Goals of this investigation

Each year, approximately 900 students apply to at least one of the EM residencies at the University of Arizona. The majority of students submit one or more SLORs with their application. Our primary objective was to describe and characterize the distribution of responses to categorical variables on the SLOR to gain an understanding of the scoring tendencies of letter writers.

METHODS

Study design and setting

This was a retrospective review of all SLORS written on behalf of all applicants to the three Emergency Medicine residency programs in the University of Arizona Health Network system in Tucson, Arizona.

The University of Arizona Health Network hosts two categorical EM residency programs (a university-based residency and a community/county hospital-based residency) and a combined EM-Pediatrics program.

Participants

All SLORs written on behalf of all the applicants to the three University of Arizona EM programs in the 2011-2012 application cycle were reviewed and included in the analysis. All candidates' applications were reviewed, and all SLORs submitted with their applications were included. SLORs were extracted from ERAS applications by the program coordinators of the University, South Campus, and EM/Pediatrics programs. Members of the study group, which included Program Directors, Associate Program Directors, Clerkship Director, and Core Medical Student Teaching Faculty, then abstracted responses for each SLOR variable. Abstraction instructions were provided by email, and spot checking of the abstraction process was conducted during data collection. Duplicate SLORs from applicants who applied to more than one of the University of Arizona EM residency programs were recorded only once. If a letter writer used more

Table 1. Variables and categories on the standard letter of recommendation with assigned scoring.

Variable	Categories	Scoring
Commitment to emergency medicine (CEM)	Outstanding	1
	Excellent	2
	Very good	3
	Good	4
Work ethic (WET)	Outstanding	1
	Excellent	2
	Very good	3
	Good	4
Development of treatment plan (DTP)	Outstanding	1
	Excellent	2
	Very good	3
	Good	4
Personality: ability to interact with others (IWO)	Superior	1
	Good	2
	Quiet	3
	Poor	4
Personality: ability to communicate with patients (CWP)	Superior	1
	Good	2
	Quiet	3
	Poor	4
Amount of guidance anticipated (GUI)	Almost none	1
	Minimal	2
	Moderate	3
Prediction of success (PRS)	Outstanding	1
	Excellent	2
	Good	3
Global assessment score (GAS)	Outstanding	1
	Excellent	2
	Very good	3
	Good	4
Likelihood of matching assessment (LOMA)	Very competitive	1
	Competitive	2
	Possible match	3
	Unlikely match	4

than one answer to a variable, for example an outstanding (scoring a 1) and an excellent (scoring a 2) for the GAS, the less favorable score was recorded for that variable on that application. Once the data collection was complete, data was de-identified by removing the applicants' ERAS number and institution and centrally collated for analysis.

Measurements and outcome variables

Anchors for each SLOR variable are listed in Table 1, along with the corresponding numerical score they were assigned in this study.

Data Analysis

Data analysis consisted of descriptive statistics of the distribution of all categorical variables collected, using Microsoft Excel for Mac 2011. The local institutional review committee approved this study.

Table 2. Descriptive analysis of variables on the standard letter of recommendation in emergency medicine (Tier 1 = highest rating, Tier 4 = lowest rating).

Variable	Rating tier 1	Rating tier 2	Rating tier 3	Rating tier 4
Commitment to emergency medicine (CEM) (n=1457)	733 (50.31%)	611 (41.94%)	104 (7.14%)	9 (0.61%)
Work ethic (WET) (n=1455)	865 (59.45%)	507 (34.85%)	77 (5.29%)	6 (0.41%)
Development of treatment plan (DTP) (n=1451)	470 (32.39%)	694 (47.83%)	270 (18.54%)	17 (1.17%)
Personality: ability to interact with others (IWO) (n=1451)	889 (61.27%)	502 (34.60%)	59 (4.06%)	1 (0.07%)
Personality: ability to communicate with patients (CWP) (n=1445)	863 (59.72%)	549 (38.00%)	33 (2.28%)	0 (0%)
Amount of guidance anticipated (GUI) (n=1448)	557 (38.47%)	797 (55.04%)	94 (6.49%)	N/A
Prediction of success (PRS) (n=1448)	776 (53.59%)	611 (42.20%)	61 (4.21%)	N/A
Global assessment score (GAS) (n=1422)	469 (32.98%)	640 (45.01%)	277 (19.48%)	36 (2.53%)
Likelihood of matching assessment (LOMA) (n=1419)	587 (41.37%)	607 (42.77%)	198 (13.95%)	27 (1.91%)

RESULTS

Characteristics of subjects

During the 2012 interview season, there were a total of 917 unique applicants with a total of 1,457 SLORs that were submitted to the three University of Arizona EM programs. Applicants had up to 4 SLORs to support their application. The average number of SLORs per applicant was 2. Twenty percent (n=184) of the total applicants did not have a SLOR included in their application. Our sample represents 26.7% of the total number of ERAS applicants for the academic year 2012.

Main results

Many of the categorical variables for these SLORs contained missing data. 2.5% of GAS scores were missing; 3.0% of LOMA scores were missing. All other variables were missing less than 1% of the time. Data from 32 applications had a variable with more than one response (<0.1% of all data). For these cases the less favorable rating was chosen.

The percentages of responses in each category for each variable are represented in the figure and Table 2. Students were placed in the top variable rating for all variables in 47% of all responses. For variables with a 4-point scale, <1% were in the lowest variable rating. One variable, CWP, had no responses in the lowest variable rating. For the two variables with 3-point scales, 6.5% (GUI) and 4.2% (PRS) of responses were in the lowest variable rating. Combined, the lowest categories were used less than 2% of the time. In total for 4-point scale variables, 91% of all responses were in one of the top two ratings. For 3-point scale variables, 94.6% of all responses were in the top two ratings.

DISCUSSION

The SLOR is an effort to standardize recommendations on behalf of medical students applying to Emergency Medicine residency programs. However, letter writers rely disproportionately on the top two categories rather than the full scale for assessment. Our findings are consistent with another recently published description of SLOR distribution of responses.⁷ There may be a number of explanations for this, including that students may only choose writers with whom they have an outstanding rapport, in effect maximizing the likelihood of an outstanding evaluation. In addition, the SLORs analyzed in this study were those submitted for emergency medicine residency applications. It is unknown how many and what the distribution of variable scores were for SLORs written but not submitted on behalf of applicants. Students frequently waive the right to see the SLOR; however they may choose not to upload a SLOR from a site where they received an unfavorable grade. Letter writers may decline writing a SLOR if they feel it will not be a favorable one for a particular student. In addition, Dean's offices have withheld SLORS with uncomplimentary categorizations. It is likely that there is a selection bias associated with our analysis in that SLORs with lower assessments were not chosen by students or their Dean's offices to support their application.

It is unclear what training, if any, on using the SLOR as a tool to evaluate and differentiate students potential evaluators receive. Although there are instructions for completing the SLOR (now SLOE) on the CORD website, it is unknown how many authors are aware of or have read the instructions. Without definitions of specific behaviors that make one

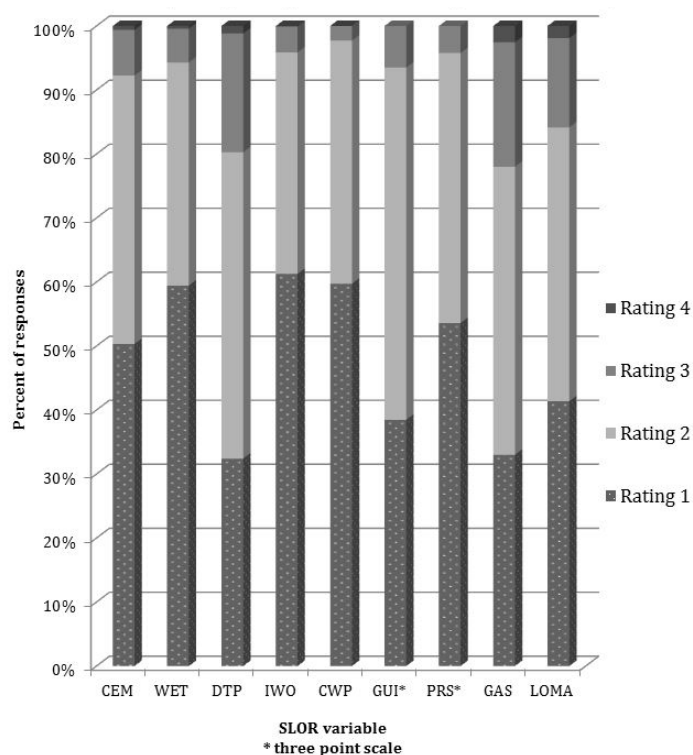


Figure. Ratings for qualifications and global assessment variables on the emergency medicine standard letter of recommendation.

student “outstanding” versus “excellent” it is left to individual SLOR writers to determine the distinction themselves. The SLOR writer’s breadth of experience working with students, as well as their experience using the SLOR as an evaluative tool limits the generalizability of the information in the SLOR when comparing students evaluated by different SLOR writers and from different medical institutions. A recent study by Beskind et. al. found that SLORs written by less experienced letter writers were more likely to have a GAS of ‘outstanding’ and a LOMA of ‘very competitive’ than more experienced letter writers.⁸

And finally, even experienced and objective letter writers may be reticent to rank a student as anything but “excellent” or “outstanding” for fear of the stigma it may carry, and potential damage to a student’s residency application.

SLOR writers were less hesitant to use lower categories on the GAS and LOMA, which are both “Global Assessment” variables rather than on the “Qualifications for EM” variables. It is possible that the qualifications variables represent a place where letter writers feel they can convey, “this is a great student, just not for our program,” rather than give the student what may be perceived as a negative letter in comparison to the other SLORs for other students in the applicant pool.

Some institutions have moved to a composite or committee SLOR. A group or departmental letter may be more objective and more likely to include the full spectrum of scaled categories for each variable. However, this may not reflect the personal experience that individual faculty have had

with an applicant. Ideally, a SLOR would accurately reflect a student’s qualifications for EM as well as a global assessment.

Recently, the CORD EM has chosen to revise the format of the SLOR to the Standardized Letter of Evaluation (SLOE). The changes made in revision of the SLOR to the SLOE reflect a greater emphasis on evaluation in addition to recommendation, and a simplification of the form in order to promote standardization across institutions. While the categories and anchors in the SLOE are very similar to those of the SLOR, it may be that our analysis does not accurately reflect the distribution of scores across the SLOE.

LIMITATIONS

This study was a cross sectional description of all SLORs written on behalf of applicants in EM at the University of Arizona, and is subject to many of the flaws of this study design. Although letter writers most commonly rank students in the highest two categories, we can only speculate as to why this occurs.

Incomplete or missing data for each variable may have affected the analysis. For example, letter writers occasionally did not rank students on one or more variables. 2.4% of GAS data and 2.6% of LOMA data were missing. Only 5 (0.3%) SLORs were missing both GAS and LOMA data. Less than 1% of all other variables were missing. We assumed this data was missing at random and simply excluded them from our analysis, rather than try to impute data. Some letter writers indicated two responses for the same variable. Data from 32 applications had a variable with more than one response. For these cases, the less favorable rating was chosen. Due to our a priori hypothesis that ratings were skewed to the more favorable side of each scale, coding these type of responses as less favorable would have had the effect of biasing our results in the opposite direction from our findings. In addition, duplicate responses were present for <0.1% of all data. Although generally low percentages of SLORs had incomplete, missing or duplicate data, this may have potentially changed the results.

While we did collect data on who wrote each SLOR and how many they wrote in the current applicant pool, we do not report that information in this analysis. It is possible that a handful of writers were responsible for a significant percentage of the total. Even though our sample size was quite large, it may be that one or two “superwriters” could skew the results because of individual tendencies in how they evaluate students. These writers could have been inclined to be more or less lenient, thereby affecting the results.

Finally, this analysis is based solely on the applicant pool of the University of Arizona EM residency programs. While the three programs received 26.7% of the total applicant pool, it is possible that this sub-population is not representative of the entire student population of interest—i.e. the total applicant pool to EM residencies – and therefore results cannot be generalized to all applicant SLORs. For example,

applicants to EM programs in the Western United States may have different SLORs than their counterparts in the Eastern or Southern regions of the country. Or, we may simply have very competitive EM residency programs, thereby limiting the number of applicants in the bottom half of the categories. Despite this, and regional differences aside, we were able to capture applicants to both a university residency program as well as community/county hospital based program representing two very common applicant pools.

CONCLUSION

SLOR letter writers were very likely to use the highest two categories for the descriptive variables when writing their letters. They were more likely to use the highest category on “Qualifications for EM” variables (CEM, WET, IWO, CWP, GUI and PRS) and to use the second highest category on Global Assessments (GAS, LOMA) and rarely used the lowest one to two categories. The lowest categories were used less than 2% of the time. Program Directors should consider the tendencies of the rating on the SLOR when reviewing SLORs of potential applicants to their programs. Although our analysis may not accurately reflect the distribution of scores across the SLOE, the categories and anchors in the SLOE are so similar to those of the SLOR the distribution of scores is unlikely to differ greatly from our findings.

Address for Correspondence: Kristi H. Grall, MD, MHPE. University of Arizona, 1501 N. Campbell Ave, Tucson, AZ 85724. Email: khgrall@aemrc.arizona.edu.

Conflicts of Interest: By the *WestJEM* article submission agreement, all authors are required to disclose all affiliations, funding sources and financial or management relationships that could be perceived as potential sources of bias. The authors disclosed none.

REFERENCES

1. Balentine J, Gaeta T, Spevack T. Evaluating applicants to emergency medicine residency programs. *J Emerg Med*. 1999;17(1):131-134.
2. Crane JT, Ferraro CM. Selection criteria for emergency medicine residency applicants. *Acad Emerg Med*. 2000;7(1):54-60.
3. 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(3):206-210.
4. Oyama LC, Kwon M, Fernandez JA, et al. Inaccuracy of the global assessment score in the emergency medicine standard letter of recommendation. *Acad Emerg Med*. 17 Suppl 2:S38-41.
5. Keim SM, Rein JA, Chisholm C, et al. A standardized letter of recommendation for residency application. *Acad Emerg Med*. 1999;6(11):1141-1146.
6. Girzadas DV Jr, Harwood RC, Dearie J, et al. A comparison of standardized and narrative letters of recommendation. *Acad Emerg Med*. 1998;5(11):1101-1104.
7. Love JN, Deiorio NM, Ronan-Bentle S, et al. SLOR Task Force. Characterization of the Council of Emergency Medicine Residency Directors' standardized letter of recommendation in 2011-2012. *Acad Emerg Med*. 2013;20(9):926-932.
8. 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*. 2013;46(6):544-550.