Are Standardized Letters of Recommendation in Residency Applications Correlated With Objective Data?

Alice L Tang[1], Javier J M Howard[1], Eshita Singh[1] iD,
Meredith E Tabangin[2], James C Wang[1], Charles M Myer[3],
Mekibib Altaye[2] and Sarah L Rohde[4]

[1]Department of Otolaryngology-Head and Neck Surgery, College of Medicine, University of Cincinnati, Cincinnati, OH, USA. [2]Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. [3]Department of Otolaryngology-Head and Neck Surgery, Cincinnati Children's Hospital Medical Center, Cincinnati OH, USA. [4]Department of Otolaryngology-Head and Neck Surgery, Vanderbilt University Medical Center, Nashville, TN, USA.

ABSTRACT

OBJECTIVE: To evaluate whether the standardized letter of recommendation (SLOR), commonly used for medical students applying to otolaryngology residency, correlates with objective data in the application.

BACKGROUND: Standardized letters of recommendation using Likert-type scales for different attributes are commonly used by evaluators because of their high interrater reliability and efficiency in preparation and interpretation. Given that these are subjectively scored, it is unknown how well these correlate with objective data.

METHODS: Applications to the University of Cincinnati otolaryngology residency were evaluated in the academic cycle of 2017-2018. Standardized letters of recommendation were scored to determine whether certain attributes were correlated with objective data (United States Medical Licensing Examination [USMLE] scores and number of presentations/publications) provided in their application. Spearman correlations were used to evaluate the strength of the relationship between the subjective score in certain attributes with objective data.

RESULTS: There were 217 applications to the University of Cincinnati that contained SLORs. Of these applications, 474 standardized letters were scored in categories of medical knowledge, research, and commitment to academic medicine. Total publications and presentations were weakly correlated with commitment to academic medicine (0.35, $P < .0001$, n = 369) and with research (0.44, $P < .0001$, n = 355). Medical knowledge was weakly correlated with Step 1 scores (0.20, $P < .0001$) and Step 2 scores (0.18, $P = .0002$).

CONCLUSIONS: Subjective research and commitment to academic medicine rating scores were weakly correlated with greater academic productivity. Similarly, medical knowledge scores were weakly correlated with Step 1/2 scores. Further research may be needed to assess how to interpret SLOR scores in addition to the information available in an otolaryngology application.

LEVEL OF EVIDENCE: 4

KEYWORDS: Standardized letters of recommendations, otolaryngology residency, LOR

Introduction

Letters of recommendation are a critical component of candidate applications for otolaryngology residency programs. They are an opportunity for programs to hear from a colleague in the field, providing additional insight into the candidate's personality, abilities, potential, communication skills, and beyond. Traditionally, narrative letters of recommendation (NLORs) are free-form, but typically contain certain elements, including the writer's relationship with the applicant, a review of the candidate's record, and an evaluation of their performance. However, studies have shown that NLORs, specifically for otolaryngology applicants, are time-intensive and do not correlate well with rank list.[1] Moreover, NLORs for multiple specialties and disciplines have been demonstrated to be vulnerable to gender bias.[1-6]

The development of the "standardized letter of recommendation" (SLOR) in 1995 by emergency residency programs sought to improve residency selection by providing consistent evaluative data for applicants and creating a time-efficient instrument and review format.[1,3,7,8] Given the common and inherent limitations with NLORs, Prager et al[9] introduced SLOR as an alternative in otolaryngology pediatric fellowship applications and found improved reliability and efficiency. This subsequently led to its use in otolaryngology residency applications where it has also been shown to save writing and reviewing time, have higher interrater reliability, and have higher correlation with ranking.[1,3]

Although the use of SLOR in otolaryngology residency is now common, there have been only a few studies that have investigated its validity.[3,10] The aim of this study was to

**Otolaryngology—Head & Neck Surgery Residency
Standardized Letter of Recommendation Form**

Applicant's Name: AAMC ERAS ID No.:

Reference Provided By:

Present Position: Email:

Institution: Telephone Number:

BACKGROUND INFORMATION

1. How long have you known the applicant?

    ☐ Less than one month   ☐ 1-6 months   ☐ 1-2 years

    ☐ 1 month   ☐ 6-12 months   ☐ 2-5 years

                                      ☐ More than 5 years

2. Nature of contact with applicant: (Check all that apply)

    ☐ Extended direct clinical contact (> 20 hrs)   ☐ Met with applicant 1-3 times to discuss application

    ☐ Extended direct research contact (>20 hrs)   ☐ Know indirectly through others/evaluations

    ☐ Limited direct clinical contact (<20 hrs)   ☐ Other (Please describe): _____

    ☐ Limited direct research contact (<20 hrs)

QUALIFICATIONS FOR OHNS: *Compare the applicant to other medical students (including those entering other specialties) by indicating with an "X" on the scales below where you rate him or her.*

1. *Patient Care* – Ability to develop and justify an appropriate differential diagnosis and a cohesive treatment plan.

    |——————————————|——————————————| ☐

    Top   Middle   Bottom   Cannot assess

2. *Medical Knowledge* – Level of general and otolaryngologic medical knowledge.

    |——————————————|——————————————| ☐

    Top   Middle   Bottom   Cannot assess

3. *Professionalism* – Quality of work ethic, altruism, professional appearance and willingness to assume responsibility.

    |——————————————|——————————————| ☐

    Top   Middle   Bottom   Cannot assess
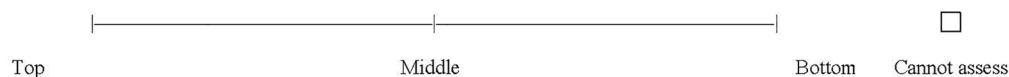
4. *Interpersonal and Communication Skills* – Ability to interact with others on the health care team and communicate in an effective and caring manner with patients.

Figure 1. Standardized letters of recommendation.

investigate whether SLOR correlated with objective data that were provided in the other components of the candidate's residency application. Although we understand that SLOR may be more efficient, it is not clear whether data provided in a subjective manner are correlated to objective information in the rest of the application.

Methods

This study was approved by the University of Cincinnati Institutional Review Board. We conducted a retrospective cohort study using all applications to the University of Cincinnati otolaryngology training program for the 2017-2018 application year. All applications with SLOR were included. The SLOR used in otolaryngology is shown in Figure 1. Applicants with more than 1 SLOR had each recommendation counted individually.

All applications containing SLORs were examined using Adobe Acrobat Pro DC (Adobe, San Jose, CA). From each SLOR, data were collected from the "Qualifications for OHNS" section under the categories of "medical knowledge," "research," and "commitment to academic medicine" by measuring the total length of each scale line and then measuring the distance of the rating from the left end of that number line of "Qualifications for OHNS." The distance measured was divided by the total length of the line to create a proportional score for each rating. Measurements were made with the Adobe measuring tool. Scaled scores closer to 0 indicate a better ranking. This method is similar to that which was used by Kominsky et al,[11] which measured the score from the right side instead of the left and then divided the score by the total length of the line to create a percentile score for each rating.

Data were also collected on each applicant's home medical school, geographic region, number and type of research publications and presentations, and United States Medical Licensing Examination (USMLE) Step 1 and Step 2 scores. We did not examine third-year clerkship grades, as there is a lack of standardization of grades across medical schools. Research publications and presentations included both peer-reviewed and non–peer-reviewed publications in medicine and nonmedicine fields alike; all publications and presentations listed in the application were counted. We did not confirm the authenticity of the research experiences listed. Descriptive statistics were calculated on all variables, including mean values with standard deviations, medians with interquartile ranges, frequencies, and proportions. Distributions of continuous variables were examined for normality using histograms and kernel densities. To assess whether objective data in the application were correlated with the subjective scores given, we used the Spearman correlation coefficient to examine the relationship between USMLE Step 1 and Step 2 board scores and the attribute "medical knowledge," as well as total publications and presentations with "research" and "commitment to academic medicine." An α of .05 was considered statistically significant. Statistical analyses were conducted with SAS version 9.4 (SAS Institute, Cary, NC).

Results

In 2017, the University of Cincinnati received 285 unique applications for otolaryngology residency. There were a total of 217 individual applicants with 474 individual SLORs included in the study. Nationwide, there were a total of 331 applicants applying for 305 spots in 2017.[12] There were 76 female and 141 male candidates in the 2017 application cycle who applied to University of Cincinnati with SLORs. The number of applicants receiving 1 SLOR was 67 (30.9%); 2 SLORs, 70 (32.3%); 3 SLORs, 56 (25.8%); and ≥4 SLORs, 24 (11%), as shown in Table 1.

Objective data were gathered from the candidate's application. The mean Step 1 score was 247.4 (SD = 12.6) and the mean Step 2 score was 254.3 (SD = 12.4). The number of publications, oral presentations, and poster presentations are listed in Table 1. Scores from the specific attributes "commitment to academic medicine," "medical knowledge," and "research" were measured, calculated, and are presented in Table 2. The number for individual scores counted may not necessarily be the total of individual SLORs of 417 given that some writers had the option of using a "not applicable" category for a particular candidate.

Scores of "medical knowledge" were weakly correlated with USMLE Step 1 scores (Spearman ρ = –0.20, *P*< .0001) and USMLE Step 2 scores (Spearman ρ = –0.18, *P*= .0002). In addition, we found that research productivity (ie, total publications and presentations) was weakly correlated with the attributes "research" (–0.44, *P*< .0001) and "commitment to academic medicine" (–0.35, *P*< .0001) (Table 3). Scatterplots are shown in Figure 2.

Table 1. Applicant demographics and characteristics.

CHARACTERISTICS	
Number of individual applicants	217
Sex, female, No. (%)	76 (35.0)
Race, No. (%)	
Asian	43 (20.7)
Black/African American	6 (2.9)
White/Caucasian	123 (59.1)
Hispanic/Latino	12 (5.8)
Other	24 (11.5)
Unknown	9 (4)
No. of evaluations per applicant, mean (SD)	2.2 (1.0)
No. of applicants with . . . No. (%)	
1 SLOR evaluation	67 (30.9)
2 SLOR evaluations	70 (32.3)
3 SLOR evaluations	56 (25.8)
≥4 SLOR evaluations	24 (11.0)
Step 1 score (n=217), mean (SD)	247.4 (12.6)
Step 2 score (n=199), mean (SD)	254.3 (12.4)
Research, median [IQR]	
No. of publications	1 [0-3]
No. of oral presentations	1 [0-3]
No. of poster presentations	3 [1-5]
Total publications and presentations	6 [4-10]

Abbreviations: IQR, interquartile range; SLOR, standardized letter of recommendation.

Table 2. Standardized letter of recommendation scores (scaled).

CTA Scaled Score (n=369)	
Median [IQR]	0.12 [0.04-0.25]
Knowledge Scaled Score (n=465)	
Median [IQR]	0.08 [0.04-0.15]
Research Scaled Score (n=355)	
Median [IQR]	0.09 [0.03-0.20]

This table reports the scaled scores for the following applicant attributes: commitment to academics, knowledge, and research. A score closer to 0 indicates better ranking as the score is calculated by measuring the mark from the left side divided by the whole line.
Abbreviations: CTA, commitment to academic medicine; IQR, interquartile range.

Discussion

In this study, our findings suggest that certain domains of the SLOR in otolaryngology are not strongly correlated with

Table 3. Correlation of Attributes to Objective Academic Activity.

SCALED SCORE—COMMITMENT TO ACADEMIC MEDICINE		
N = 376	SPEARMAN ρ	P VALUE
Total publications and presentations	−0.35	<.0001
Total presentations (oral and poster)	−0.27	<.0001
No. of publications	−0.37	<.0001
No. of poster presentations	−0.26	<.0001
No. of oral presentations	−0.12	.03
SCALED SCORE—RESEARCH		
N = 358	SPEARMAN'S ρ	P-VALUE
Total publications and presentations	−0.44	<.0001
Total presentations (oral and poster)	−0.39	<.0001
No. of publications	−0.33	<.0001
No. of poster presentations	−0.34	<.0001
No. of oral presentations	−0.28	<.0001

This table reports the correlation of the scaled scores of commitment to academic medicine and research attributes to objective academic activity. A $P < .05$ was considered significant for the analysis.

objective data provided by a candidate's application. Although we anticipated that the number of publications and presentations in an applicant's experience would be associated with the categories of "commitment to academic medicine" and "research" in the SLOR, we did not see a strong correlation. Similarly, "medical knowledge" was weakly correlated with USMLE Step 1 scores.

In 2017, our institution received 474 SLORs on behalf of medical students applying to otolaryngology residency for that year. The use of SLOR has grown since the cycle of 2011-2012 when it was first introduced to otolaryngology by the Otolaryngology Program Directors Organization.[1] Following its introduction, few studies have evaluated its validity in selecting medical students for otolaryngology residency.[1-3,11,13,14] Kimple et al[3] found that only certain attributes of the SLOR were correlated with successful matching in the field. Interestingly, attributes of "professionalism," "procedural skills," "research," "commitment to otolaryngology," and "commitment to academic medicine" were not correlated with matching into the otolaryngology specialty.

The benefits of SLOR was surveyed among program directors and chairs after its introduction, and respondents thought that, indeed, using SLOR was more efficient and improved the communication of relevant and useful applicant information.[13] As compared with NLOR, the use of SLOR in otolaryngology and other medical fields has demonstrated an improvement in efficiency in preparing and
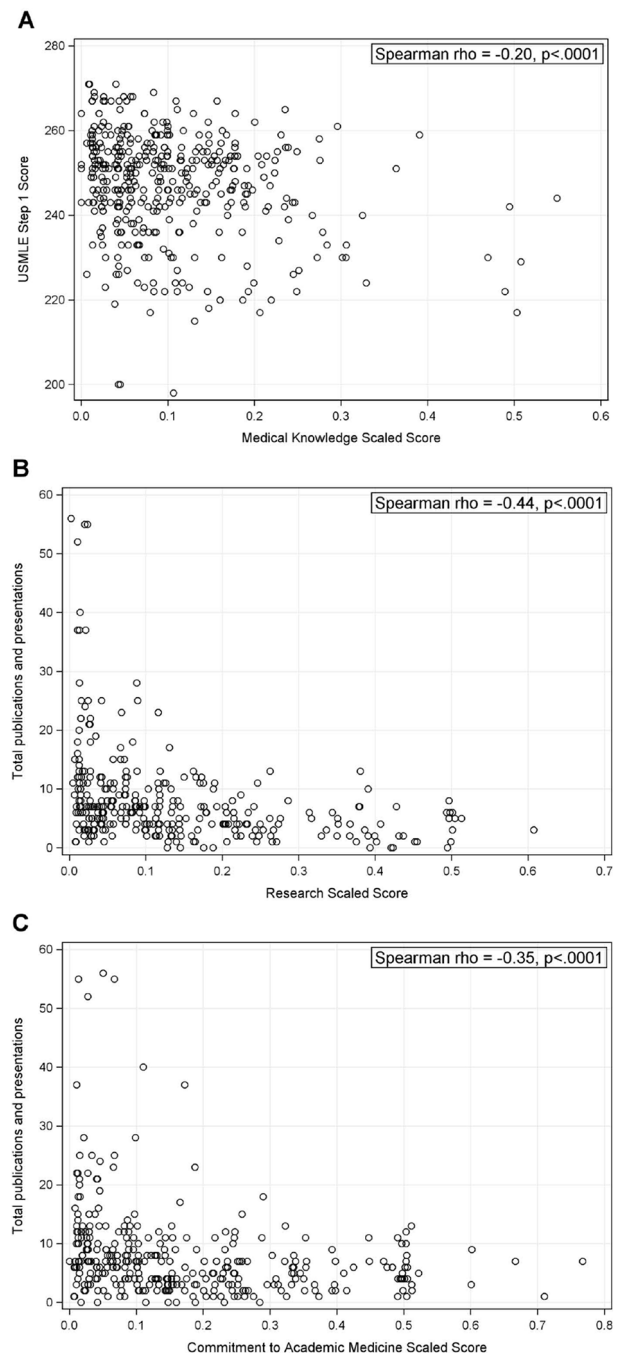
Figure 2. Scatterplots of (A) USMLE Step 1 score by Medical Knowledge Scaled Score, (B) total publications and presentations by Research Scaled Score, and (C) total publications and presentations by Commitment to Academic Medicine Scaled Score.

interpreting recommendations and has higher interrater reliability.[1,15] However, one criticism of the SLOR is that the narrative component was found to unfortunately demonstrate similar implicit gender bias problems seen in traditional NLOR, suggesting that future iterations of the SLOR consider abandoning the narrative option.[2,14] Young[16] found that the impression of an applicant drawn from either NLOR or SLOR and a face-to-face interaction had moderate but not excellent concordance.

In an effort to improve the selection method for otolaryngology residency, we focused on the use and validity of the SLOR in this process. We specifically selected the attributes ("research," "commitment to academic medicine," and "medical knowledge") because these specific domains are theoretically objectively represented in other parts of the application (ie, publications, presentations, participation in otolaryngology projects, USMLE Step scores, etc). Our data support that these domains, some of which are already not well correlated with successful matching,[3] are also not correlated with other quantitative measures.

However, this observation could imply that scores given through the SLOR may be providing additional information lost in the interpretation of the objective data. For instance, some applicants may have extensive, meaningful research experiences that are not accurately represented by a raw number of presentation and publication output. An evaluator who has witnessed strong commitment to a project and found that the student excelled in the follow-through with research tasks may feel that the applicant deserves a better score in "commitment to academic medicine" and "research," despite a lower number of publications or presentations. Currently, the selection for medical training can overemphasize *quantity over quality*,[17] and further research is required to evaluate whether subjective SLOR grades are providing some facet of *quality*.

Refining selection criteria for potential trainees in residency is greatly needed. In smaller and more competitive fields, when there are more applicants than spots available, academic ability and medical knowledge are frequently used as distinguishing features in selecting between applicants. Thus, the "medical knowledge" attribute on the SLOR can be useful. The lack of correlation between "medical knowledge" scores and USMLE scores can be explained by the fact that USMLE Step exams capture a very limited scope of an applicant's medical knowledge, as student performance may vary based on other factors such as the student's test-taking abilities, educational background, and simply the content tested on that particular exam date. Furthermore, "medical knowledge" may be capturing performance from an extended time period of a clinical rotation, as opposed to one point in time on the day of the USMLE. The "medical knowledge" category might be interpreted by evaluators as reflective of the student's mastery of otolaryngology medical knowledge in a more practical setting, like the clinic or operating room. This could be an opportunity to revise and define the "medical knowledge" attribute of the SLOR to describe "otolaryngology fund of knowledge and preparedness," which can potentially be useful for application reviewers and is an area that may not be reflected elsewhere in the application.

In contrast, the lack of correlation can also be because the examiner may not be able to give an accurate assessment of "medical knowledge" due to limited contact with students and depend on the performance evaluations of others—calling into question the accuracy and reliability of a subjective score for this attribute.[11] Ultimately, we must consider what additional information is being added by a subjective score in "medical knowledge" after reviewing the global picture of the applicant with grades and Step 1/2 performances also accounted for.

There are limitations in this study. As the distance was calculated manually on Adobe, there is potential for human error and interuser variability. Because most applicants' scores were concentrated in the left-most region of the scale (highest score), small errors in calculating distance could result in significant differences in scaled score. Furthermore, a significant portion of the research output listed on applications is unable to be verified, and some may represent inaccurate or "ghost" publications.[18] Further investigations could include performing this data analysis over several application cycles to improve accuracy—as this study only looked at 1 application cycle at 1 institution. In addition, as mentioned previously, we did not examine third-year clerkship grades as an objective measure given the limited standardization across schools. However, future studies could include this as a factor of "medical knowledge." Our data found a weak correlation between "medical knowledge" and USMLE Step 1 scores, perhaps due to the overall small variation in ratings across the 10 categories of the SLOR; Kominsky et al[11] previously reported that ratings in each of the 10 categories clustered within the top 2 deciles.

Despite our analysis and criticism of these specific attributes, we are strong advocates of using the SLOR to convey and quantify more immeasurable qualities of an applicant: interpersonal communication skills, initiative and drive, professionalism, and procedural skills. These domains from a letter of support can communicate an assessment of a candidate's potential when numbers and past experiences cannot. In addition, the attributes of "research," "commitment to academic medicine" and "medical knowledge," although not well correlated with objective data in the application, could be used in a different perspective on the applicant of these attributes. We also appreciate through a review of the literature that the use of SLOR can possibly reduce implicit gender bias in its preparation.[14] Further research is needed to continue to refine our efforts to guide residency programs in selecting successful applicants.

Conclusions

Standardized letters of recommendation in otolaryngology residency applications are efficient in preparation and interpretation; however, reviewers should consider how to interpret this additional information. "Medical knowledge," "commitment to academic medicine," and "research" are weakly correlated to objective measures provided in a candidate's application; however, this does not imply that these scores have no value. Future research is needed to explore whether these subjective scores may be providing additional insight for these categories.

Author Contributions

ALT: study concept and design, data collection, statistics, manuscript writing, manuscript editing

JJMH: data collection, statistics, manuscript writing, manuscript editing

ES: data collection, statistics, manuscript writing, manuscript editing

MET: statistics, manuscript writing, manuscript editing

JCW: data collection, manuscript editing

CMM: manuscript editing

MA: statistics, manuscript writing, manuscript editing

SLR: study concept and design, manuscript editing

ORCID iD

Eshita Singh https://orcid.org/0000-0003-0033-736X

REFERENCES

1. Perkins JN, Liang C, McFann K, Abaza M, Streubel SO, Prager J. Standardized letter of recommendation for otolaryngology residency selection. *Laryngoscope*. 2013;123:123-133.
2. 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;118:1335-1344.
3. Kimple AJ, McClurg SW, Del Signore AG, Tomoum MO, Lin FC, Senior BA. Standardized letters of recommendation and successful match into otolaryngology. *Laryngoscope*. 2016;126:1071-1076.
4. Turrentine FE, Dreisbach CN, St Ivany AR, Hanks JB, Schroen AT. Influence of gender on surgical residency applicants' recommendation letters. *J Am Coll Surg*. 2019;228:356-365.
5. Li S, Fant AL, McCarthy DM, Miller D, Craig J, Kontrick A. Gender differences in language of standardized letter of evaluation narratives for emergency medicine residency applicants. *AEM Educ Train*. 2017;1:334-339.
6. Trix F, Psenka C. Exploring the color of glass: letters of recommendation for female and male medical faculty. *Discourse Soc*. 2003;14:191-220.
7. Keim SM, Rein JA, Chisholm C, et al. A standardized letter of recommendation for residency application. *Acad Emerg Med*. 1999;6:1141-1146.
8. 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;20:926-932.
9. Prager JD, Perkins JN, McFann K, Myer CM III, Pensak ML, Chan KH. Standardized letter of recommendation for pediatric fellowship selection. *Laryngoscope*. 2012;122:415-424.
10. Hu AC, Gu JT, Wong BJF. Objective measures and the standardized letter of recommendation in the otolaryngology residency match [published online ahead of print May 7, 2019]. *Laryngoscope*. doi:10.1002/lary.28054.
11. Kominsky AH, Bryson PC, Benninger MS, Tierney WS. Variability of ratings in the otolaryngology standardized letter of recommendation. *Otolaryngol Head Neck Surg*. 2016;154:287-293.
12. National Resident Matching Program. Results and data: 2017 main residency match. https://www.nrmp.org/wp-content/uploads/2017/06/Main-Match-Results-and-Data-2017.pdf. Updated 2017. Accessed November 13, 2019.
13. Messner A, Teng M, Shimahara E, et al. A case for the standardized letter of recommendation in otolaryngology residency selection. *Laryngoscope*. 2014; 124:2-3.
14. Friedman R, Fang CH, Hasbun J, et al. Use of standardized letters of recommendation for otolaryngology head and neck surgery residency and the impact of gender. *Laryngoscope*. 2017;127:2738-2745.
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. Young VN. Letters of recommendation: association with interviewers' perceptions and preferences. *Otolaryngol Head Neck Surg*. 2017;156:1108-1113.
17. Patterson F, Knight A, Dowell J, Nicholson S, Cousans F, Cleland J. How effective are selection methods in medical education? a systematic review. *Med Educ*. 2016;50:36-60.
18. Ishman SL, Smith DF, Skinner ML, et al. Unverifiable publications in otolaryngology residency applications. *Otolaryngol Head Neck Surg*. 2012;147:249-255.