

Journal of Contemporary Medical Education

available at www.scopemed.org

Original Research

USMLE step one, otolaryngology in-training examination scores, and the American Board of Otolaryngology examination performance

Beth N. McNulty,¹ Jeffrey M. Bumpous,² Jarrod A. Little,³ Julie L. Goldman⁴

¹Resident Division of Otolaryngology-Head and Neck Surgery; ²J. Samuel Baumgardner Endowed Professor and Chief, Division of Otolaryngology-Head and Neck Surgery; ³Division of Plastic and Reconstructive Surgery and Otolaryngology; ⁴Division of Otolaryngology, Department of Surgery, University of Louisville

Received: November 29, 2012

Accepted: January 16, 2013

Published Online: March 05, 2013

DOI: 10.5455/jcme.20130110113516

Corresponding Author:

Beth Nicole McNulty,
University of Louisville School of
Medicine, Louisville, KY, USA
beth.payne@louisville.edu

Keywords: USMLE Step One, in-training examination, board performance

ABSTRACT

To determine if U. S. Medical Licensure Examination (USMLE) Step One scores have a positive correlation with performance on the annual Otolaryngology Training Examination (OTE) and subsequent performance on the American Board of Otolaryngology (ABOto) examination. Study Design: A retrospective cohort study of 70 residents completing Otolaryngology residencies from 2005-2009. 103 otolaryngology training programs were contacted and USMLE Step 1 scores, OTE scores, and performance on the ABOto examination were collected for the graduating chief residents from 2005-2009 from a limited number of residency programs. The associations between the examinations were evaluated using Spearman's rank correlation coefficient and Pearson's correlation. USMLE Step 1 scores and board performance were provided for 70 residents from 7 Otolaryngology residency programs. Positive correlations were found using Pearson's and Spearman's correlation to be significant for USMLE Step 1 scores compared with scores on the OTE-2, OTE-3, OTE-4, OTE-5 ($p < 0.05$) and the board examination ($p < 0.05$). Statistically significant correlation was seen between the OTE-4 score and performance on the board examination using both Pearson's and Spearman's ($p < 0.05$), but not for the OTE-2, OTE-3, OTE-5 and performance on the written board examination. An objective criterion for interview selection is the USMLE Step 1 score; we found a positive correlation with performance on the annual OTE and subsequent performance on the American Board of Otolaryngology examination. Only the OTE-4 score was found to positively correlate with board examination performance. Larger sample size and longitudinal follow-up is necessary to establish the significance of these findings.

© 2013 GESDAV

INTRODUCTION

The field of otolaryngology is increasingly competitive among medical student applicants. According to the National Residency Match Program (NRMP) 2009 data, dermatology, neurological surgery, orthopedic surgery, and otolaryngology were the most competitive specialties for medical school seniors [1]. In 2007, 266 residency positions were offered in otolaryngology and there were 353 applicants, or 1.33 applicants/ position [2]. The number of residency positions increased to 283 in 2011, but the number of applicants has also

increased to 385, for a slight overall ratio increase of 1.36 applicants/position [3]. In a competitive specialty with an increasing number of applicants, what factors may be used as predictors of success in an otolaryngology residency?

In 1990, an investigation of residency applicant characteristics associated with successful matching in Otolaryngology showed a correlation between excellent academic performance in medical school and residency

performance in otolaryngology [4]. Excellent academic performance was defined as a student meeting at least three of the four following criteria: (1) grade point average (GPA) >3.4, honors on junior surgery and medicine rotations, (2) National Board of Medical Examiners (NBME) score higher than 650, (3) class rank >85th percentile, and (4) election to Alpha Omega Alpha (AOA) [4]. A 2006 study found that AOA election, an exceptional trait mentioned in a letter of recommendation, United States Medical Licensure Examination (USMLE) scores, and interviewer impressions were predictors for success among residents trained in an academic otolaryngology program [5]. Multiple studies report that letters of recommendation (including Dean's letters) [6,7], medical school grades [8], and preclinical and clinical faculty evaluations [9,10,11,12] do not predict residency performance [6]. Many programs use USMLE Step 1 score minimums as one means of narrowing the applicant pool. The USMLE Step 1 examination is designed to measure a student's understanding of important basic science concepts and the ability to apply that knowledge to the practice of medicine. Most medical students take the examination after the completion of their sophomore year. To be eligible for the examination, one must either be: 1. A medical student officially enrolled in, or a graduate of, a US or Canadian medical school program leading to the Medical Doctor (MD) degree that is accredited by the Liaison Committee on Medical Education; 2. A medical student officially enrolled in, or a graduate of, a US medical school leading to the Doctor of Osteopathic medicine (DO) degree that is accredited by the American Osteopathic Association; 3. A medical student officially enrolled in, or a graduate of, a medical school outside the United States and Canada who meets the eligibility criteria of the Educational Commission for Foreign Medical Graduates [13]. The USMLE Step 1 score is the only quantitative factor common to all applicants and educational institutions [14]. Very little information has been reported with regard to the use of the USMLE Step 1 score specifically as a means to predict resident success in otolaryngology. Therefore, the objective of this study was to determine if USMLE Step 1 scores have a positive correlation with performance on the annual Otolaryngology Training Examination (OTE) and subsequent performance on the American Board of Otolaryngology (ABOto) written examination. Secondary endpoints included determining if OTE

performance stratified by resident year predicted performance on the ABOto certifying examination.

METHODS

The study was granted IRB approval by the University of Louisville Human Subjects Protection Program. The American Medical Association- Fellowship and Residency Electronic Interactive Database Access (FREIDA) was used to provide the contact information for 103 otolaryngology residency training programs in the United States. The program directors were contacted via email requesting the following de-identified information for the graduating chief residents from 2005-2009: USMLE Step 1 score, OTE overall scaled score for each year of residency, and the ABOto written examination performance (pass/fail). Three electronic mail requests were made and if a program showed initial interest in participation but failed to provide the requested data, a telephone call to the program coordinator or director was made. Indirect linkages between residency program and scores were destroyed after data collection. Board pass rates for the ABOto written certification examination for the years 2005-2009 were obtained from the American Board of Otolaryngology. The associations between the examinations were evaluated using Pearson's correlation and Spearman's rank correlation coefficient.

RESULTS

Twelve of 103 program directors responded, for an initial response rate of 11.6%. Seven programs geographically distributed across the United States agreed to participate and provided the requested data for a participation rate of 6.7%. One program provided Step 1 scores and board performance but did not provide the scaled scores. Another program could not provide complete resident data. USMLE Step 1 scores and the ABOto written examination performance were provided for 70 residents. OTE 1 scores were only available for four residents, as many programs do not require their first year residents participate in the OTE; these scores were therefore not considered for statistical analysis. OTE 2 scores were provided for 53 residents and OTE 3-5 scores for 55 residents.

Figure 1 shows USMLE Step 1 scores for all residents included in the study; the USMLE Step 1 score mean was 232.9, median 236, maximum 271, and minimum

was 184. ABOto written examination performance consisted of 4/70 failures, for a 94.3% pass rate. The USMLE Step 1 scores of the failing residents were 184, 197, 218, 237 (mean 232). The OTE-5 scaled scores for the failing residents were 5.71, 6.57, 6.52, and 6.27, with a mean of 6.57 being equivalent to the overall mean for the OTE-5. The OTE-4 scores for the failing residents were 6.66, 6.25, 6.22, and 5.71, with a mean of 6.21, slightly lower than the overall mean of the OTE-4 of 6.58 (see Table 1). Three-fourths of the surveyed residents passed the written certifying examination within three attempts. One quarter failed two attempts and did not make a third attempt. The written examination pass rates for the ABOto for recent years: 2005 - 95%, 2006 - 93%, 2007 - 88%, 2008 - 86%, 2009 - 88%, 2010 - 89%.

Positive correlations were found using Pearson's and Spearman's correlation to be significant for USMLE Step 1 scores compared with scores on the OTE-2, OTE-3, OTE-4, OTE-5, and the board examination ($p < 0.05$) (see Table 2). Statistically significant correlation was seen between the OTE-4 score and

performance on the board examination using both Pearson's ($p < 0.039$) and Spearman's ($p < 0.049$), but not for the OTE-2, OTE-3, OTE-5 and performance on the ABOto written certification examination (see Table 3, 4).

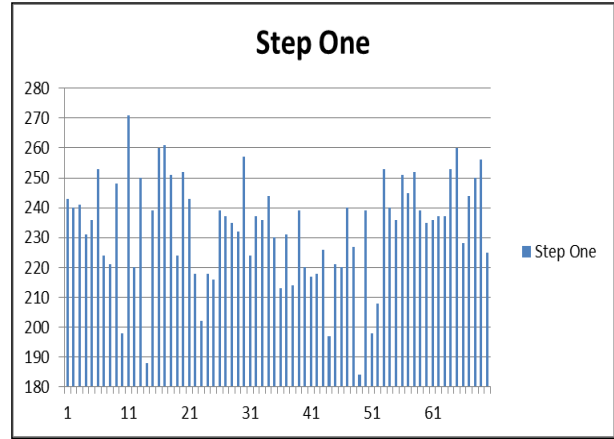


Figure 1. USMLE Step 1 scores for residents 1-70.

Table 1. Overall scaled scores for OTE 2-5.

Test	N =	Mean	Median	Maximum	Minimum
OTE 2	53	5.92	5.96	6.98	4.98
OTE 3	55	6.37	6.36	7.18	5.6
OTE 4	55	6.58	6.63	7.57	5.71
OTE 5	55	6.57	6.64	7.47	3.62

Table 2. Correlation of USMLE Step 1 scores. **Correlation is significant at the $p < 0.01$ level (2-tailed). *Correlation is significant at the $p < 0.05$ level (2-tailed).

		Step One	OTE 2	OTE 3	OTE 4	OTE 5	Boards
Spearman's Rho	Correlation Coefficient	1.000	.562**	.380**	.561**	.368**	.264*
	Sig. (2-tailed)	.	.000	.004	.000	.006	.027
	N	70	53	55	55	55	70
Pearson's	Pearson Correlation	1	.511**	.370**	.537**	.272*	.331**
	Sig. (2-tailed)		.000	.005	.000	.045	.005
	N	70	53	55	55	55	70

Table 3. Pearson correlation for OTE scores and performance on the ABOto written examination. **Correlation is significant at the p< 0.01 level (2-tailed). *Correlation is significant at the p<0.05 level (2 tailed).

		Step One	OTE 2	OTE 3	OTE 4	OTE 5	Boards
OTE 2	Pearson Correlation	.511**	1	.599**	.704**	.149	.211
	Sig. (2-tailed)	.000		.000	.000	.286	.129
	N	53	53	53	53	53	53
OTE 3	Pearson Correlation	.370**	.599**	1	.621**	.101	.220
	Sig. (2-tailed)	.005	.000		.000	.465	.106
	N	55	53	55	55	55	55
OTE 4	Pearson Correlation	.537**	.704**	.621**	1	.210	.279*
	Sig. (2-tailed)	.000	.000	.000		.123	.039
	N	55	53	55	55	55	55
OTE 5	Pearson Correlation	.272*	.149	.101	.210	1	.156
	Sig. (2-tailed)	.045	.286	.465	.123		.257
	N	55	53	55	55	55	55
Boards	Pearson Correlation	.331**	.211	.220	.279*	.156	1
	Sig. (2-tailed)	.005	.129	.106	.039	.257	
	N	70	53	55	55	55	71

Table 4. Spearman's Correlation for the OTE scores and performance on the ABOto written certification examination. **Correlation is significant at the p< 0.01 level (2-tailed). *Correlation is significant at the p<0.05 level (2 tailed).

		Step One	OTE 2	OTE 3	OTE 4	OTE 5	Boards	
OTE 2	Correlation Coefficient	.562**	1.000	.668**	.700**	.395**	.184	
	Sig. (2-tailed)	.000	.	.000	.000	.003	.186	
	N	53	53	53	53	53	53	
OTE 3	Correlation Coefficient	.380**	.668**	1.000	.577**	.391**	.205	
	Sig. (2-tailed)	.004	.000	.	.000	.003	.133	
	N	55	53	55	55	55	55	
Spearman's rho	OTE 4	Correlation Coefficient	.561**	.700**	.577**	1.000	.320*	.267*
	Sig. (2-tailed)	.000	.000	.000	.	.017	.049	
	N	55	53	55	55	55	55	
OTE 5	Correlation Coefficient	.368**	.395**	.391**	.320*	1.000	.252	
	Sig. (2-tailed)	.006	.003	.003	.017	.	.064	
	N	55	53	55	55	55	55	
Boards	Correlation Coefficient	.264*	.184	.205	.267*	.252	1.000	
	Sig. (2-tailed)	.027	.186	.133	.049	.064	.	
	N	70	53	55	55	55	71	

DISCUSSION

Determination of accurate predictors of success in residency and for passing national board certification examinations have been repeatedly explored by medical specialties including orthopedic surgery, radiology, and obstetrics and gynecology. A study of obstetrics and gynecology residents found a positive correlation between USMLE scores and Council on Resident Education in Obstetrics and Gynecology (CREOG) in-training examination scores for 24 residents at a single institution [15]. Orthopedic surgery programs have reported conflicting results. A study by Crawford et al showed that residents who scored >220 on the USMLE Step 1 were more likely to pass the American Board of Orthopedic Surgery (ABOS) part I exam (ABOS-I), while residents who scored less than 202 were more likely to fail. Orthopedic in-training examination (OITE) percentile rank scores, particularly those from the PGY-3 year, provide a more accurate prediction of passage of parts I and II of the ABOS certifying examination [16].

Daly et al measured success in otolaryngology residency using OTE scores, teaching or research awards, presentations at national meetings, publications in peer reviewed journals, and completion of masters of science or doctoral degree in otolaryngology [6]. This assumes standardization of OTE scores; however, according to ABOto administration, the OTE scaled score cannot be used to compare performance between examination years. The OTE Scale score is a transformation of the raw score and the data are rescaled each year. For example, a 5.50 from the 2009 training examination does not represent the same level of proficiency as a 5.50 from the 2011 training examination. The ABOto advises against the use of the OTE as the sole measure of resident success, instead encouraging program directors to use the aggregate performance of their residents when evaluating the strengths and weaknesses of their educational program [17]. Anecdotal evidence would suggest that most programs use the OTE score as a component of evaluation of overall resident performance, primarily because it is one of the only objective criteria available for use and common to all residents, with an understanding that this is not a sufficient sole measure of resident performance.

Our results suggest that the USMLE Step 1 score may be an objective criterion for otolaryngology residency applicant selection. A positive correlation was seen between USMLE Step 1 scores and performance on the annual OTE and subsequently on the American Board of Otolaryngology written certifying examination. Somewhat surprisingly, the OTE-4 score but not the OTE-5 score was found to positively correlate with board examination performance. Similar studies have

attributed this to increased work load and responsibility placed on chief residents, impacting their preparation time for the examination [16]. One could also surmise a plateau effect or overconfidence based on previous exam performance as residents near the completion of their training.

The USMLE Step 1 examination is not designed to predict the future performance of physicians [15]. In using USMLE Step 1 minimums, programs could be placing excessive emphasis on selecting residents with excellent test taking skills that improve their chances of success on the written board examination, to the exclusion of appropriate evaluation of future clinical skills. So what does performance on the USMLE Step 1, OTE, or written board examination tell us about the clinical skills of a residency applicant or resident? The USMLE Step 2 exam is designed to assess clinical knowledge and possibly would better predict clinical performance in resident applicants. At this point, there are no foolproof identifiers with which to screen resident applicants and identify those who will excel. We continue to attempt to refine the process and make it less subjective, which would be helpful to both the applicant and program.

CONCLUSION

One objective criterion used by otolaryngology residency programs to narrow the applicant pool for interview selection is the USMLE Step 1 score. These scores were found to have a positive correlation with performance on the annual OTE and subsequent performance on the American Board of Otolaryngology written certifying examination. While the USMLE Step 1 examination may be used to predict future examination performance, it was not designed to predict future clinical performance and using it in this manner is inconsistent with its original intent [14].

No correlation was seen between OTE-2, 3, or 5 scores and board passage rates. The OTE-4 score was found to positively correlate with subsequent board examination performance. Programs may consider using the OTE-4 score to identify residents at risk for poor performance on the written board examination. The OTE should not be used as the sole measure of resident success but as an objective component contributing to the evaluation of the resident's overall performance. A follow-up survey to Program Directors regarding their use of the OTE scores would be an interesting addition to this study.

Larger numbers of responses and longitudinal follow-up are necessary to establish the significance of our findings. A larger data set, including the majority of

residency programs in Otolaryngology would produce a more robust data set for analysis and could potentially further confirm these limited results.

Acknowledgements

We would like to thank Arun K. Gadre MD and Swapna K. Chandran MD for their support of this project. We would further like to acknowledge Dr. Robert Miller of the American Board of Otolaryngology for providing information on the intent and limitations of the ABOto OTE. Statistical assistance was provided by Savitri N. Appana, Biostatistician, Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville.

REFERENCES

1. National Residency Matching Program. Medical Schools, Students, Teaching Hospitals Celebrate Biggest "Match Day" on Record. Available via <https://www.aamc.org/newsroom/newsreleases/2009/91932/090319.html> (Accessed 25 August 2010).
2. National Resident Matching Program and Association of American Medical Colleges. Charting Outcomes in the Match Characteristics of Applicants Who Matched to Their Preferred Specialty in the 2007 NRMP Main Residency Match, 2nd Edition. August 2007. Available via <http://www.nrmp.org/data/chartingoutcomes2007.pdf> (Accessed 24 March 2012).
3. National Resident Matching Program and Association of American Medical Colleges. Charting Outcomes in the Match Characteristics of Applicants Who Matched to Their Preferred Specialty in the 2011 Main Residency Match, 4th Edition. August 2011. Available via <http://www.nrmp.org/data/chartingoutcomes2011.pdf> (Accessed 24 March 2012).
4. Calhoun KH, Martinez SA, Stevens MH, Hokanson JA, Bailey BJ. The resident selection process in otolaryngology-head and neck surgery. *Arch Otolaryngol Head Neck Surg* 1990;116(9):1041-43.
5. Daly KA, Levine SC, Adams GL. Predictors for resident success in otolaryngology. *Journal of the American College of Surgeons* 2006;202(4):649-54.
6. Clemente M, Michener WM. The dean's letter of recommendation and internship performance. *J Med Educ* 1976;51:590-92.
7. Greenberg AG, Doyle J, McClure DK. Letters of recommendation for surgical residencies: what they say and what they mean. *J Surg Res* 1994;56:192-98.
8. Korman M, Stabtefield RL. Medical school evaluation and internship performance. *J Med Educ* 1971;46:670-73.
9. Erlandson EE, Calhoun JG, Barrack FM, et al. Resident selection: applicant selection criteria compared with performance. *Surgery* 1982;92:270-75.
10. Gunzburger LK, Frazier RG, Yang LM, Rainey ML, Wronski T. Premedical and medical school performance in predicting first year residency performance. *J Med Educ* 1987;62(5):379-84.
11. Markert RL. Cognitive predictors of residency performance. *Res Staff Phys* 1989;Aug:53-58.
12. Veloski J, Herman MW, Gonnella JS, Zeleznik C, Kellow WF. Relationships between performance in medical school and first postgraduate year. *J Med Educ* 1979;54:909-16.
13. USMLE 2013 Bulletin of Information. Available via <http://www.usmle.org/pdfs/bulletin/2013bulletin.pdf> (Accessed 6 January 2013).
14. Jolly P. Charting Outcomes in the Match: Characteristics of Applicants Who Matched to Their Preferred Specialty in the 2005 NRMP Main Residency Match. July 2006. Available via <http://www.nrmp.org/data/matchoutcomes2006.pdf> (Accessed 27 August 2010).
15. Armstrong A, Alvero R, Nielsen P, Deering S, Robinson R, Frattarelli J, Sarber K, Duff P, Ernest J. Do U.S. medical licensure examination step 1 scores correlate with council on resident education in obstetrics and gynecology in-training examination scores and American board of obstetrics and gynecology written examination performance? *Mil Med* 2007;172(6):640-43.
16. Crawford CH, Nyland J, Roberts CS, Johnson JR. Relationship Among United States Medical Licensing Step I, Orthopedic In-Training, Subjective Clinical Performance Evaluations, and American Board of Orthopedic Surgery Examination Scores: A 12-Year Review of an Orthopedic Surgery Residency Program. *Jour of Surg Ed* 2010;67(2):71-78.
17. American Board of Otolaryngology 2011 Otolaryngology Training Examination Understanding your score report. Available via <https://www.aboto.org/2011%20Understanding%20Your%20Score%20Report.pdf> (Accessed 9 May 2011).

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.