The Laryngoscope
Lippincott Williams & Wilkins, Inc.

Program Directors' Opinions about Surgical Competency in Otolaryngology Residents

M. M. Carr, DDS, MD, MEd, FRCSC

Objectives: **The purpose of this study was to determine whether certain surgical procedures could be used as benchmark skills to monitor resident progress in developing surgical competency.** ***Study Design:*** **Survey.** ***Methods:*** **A two-stage survey was sent to otolaryngology residency program directors in the United States. Respondents were given a list of otolaryngology surgical procedures monitored by the American Board of Otolaryngology (ABO) and were asked to indicate whether they felt residents should be able to do each as a primary surgeon. The appropriate level of training for competency in each procedure and estimated number of procedures to competency was indicated by respondents.** ***Results:*** **Respondents selected 16 common procedures they felt residents at different levels of training should be able to perform independently. There were discrepancies between estimated number of procedures needed for competence and the numbers reported by ABO graduates.** ***Conclusions:*** **Surgical skill is one aspect of clinical competency, and this indicates agreement among program directors with regard to a set of benchmark skills we can use for concentrated evaluation efforts.** ***Key Words:*** **Surgical education, residency, surgical skills.**

Laryngoscope, **115:1208–1211, 2005**

INTRODUCTION

There are numerous surgical procedures within the scope of otolaryngology. Some form the core of the specialty, and these are useful to evaluate how well a program prepares residents for practice and how well a resident has mastered the curriculum. There is no published information on which procedures should make up this core group.

Abundant attention has recently been put on training for procedural skills, especially in surgery.[1,2] Models, both high and low fidelity, simulation apparati, and virtual reality systems are being developed in the hope that bench training translates to the operating room.[3,4] Valid and reliable evaluation of real-life skills in the operating room is still in its infancy, and there is good evidence that residents are not good self-assessors in this domain.[5]

Developing and testing models and evaluation schemata is time consuming and expensive. Because there are almost 100 surgical procedures comprising the otolaryngology curriculum, being able to rationally pick out skills that are widely performed in adequate numbers and that represent each level of resident mastery would contribute to making the process of developing teaching and assessment programs more efficient.

Purpose of Surveys

1. To identify which surgical procedures are appropriate to use as benchmarks for surgical skill in otolaryngology residents at each level of training.

2. To identify the number of cases residency supervisors feel are necessary to achieve competency in these surgical skills and compare them with actual numbers reported by otolaryngology graduates.

MATERIALS AND METHODS

This was a survey study undertaken following a modified Delphi technique and Dillman method.[6] The survey is based on the American Board of Otolaryngology (ABO) case reporting list, which is used to record a resident's surgical experience. All procedures on this list where the average number performed as the primary surgeon by graduating otolaryngology residents exceeded two were included on this survey, resulting in 73 procedures appearing on the survey. The survey was mailed to all U.S. otolaryngology residency program directors (n = 102). They were asked to judge whether they felt that a procedure was masterable at the resident level and if so, at what level of training. They were asked to estimate the number of cases as primary surgeon that would be needed to achieve competence in each procedure under typical circumstances. Surveys were returned by fax. A second mailing was sent only to nonresponders.

Data were summarized by frequencies and compared with chi-square tests using SPSS 12.0 (SPSS Inc., Chicago IL). A second survey was constructed listing the most frequently cited procedures for each level of training from postgraduate year 1 to 5 (PGY1–PGY5), and the same group was asked to indicate how strongly they agreed with the results. This survey contained ABO average case numbers for graduating otolaryngology residents for each procedure listed. Again, a second mailing was sent to nonresponders.

From the Department of Surgery, Penn State College of Medicine, Hershey, Pennsylvania, U.S.A.

Editor's Note: This Manuscript was accepted for publication February 17, 2005.

Send Correspondence to Dr. M. M. Carr, Department of Surgery, Penn State College of Medicine, 500 University Drive, PO Box 850, Mail Code H091, Hershey, PA 17033, U.S.A.

DOI: 10.1097/01.MLG.0000163101.12933.74

Data were again summarized and compared with nonparametric statistics. This study was approved by the local institutional review board (protocol #18424).

RESULTS

Survey 1

The response rate for survey 1 was 62%. All respondents completed, at least in part, the questions asking them to estimate the level of training at which residents usually can perform each listed procedure competently. However, only 45 of the 63 respondents (44% of the total group) responded to the question asking them to estimate the number of cases a trainee would usually need to do to be able to do the specific procedure competently. Two respondents indicated on the survey that they felt this number was dependent on the individual resident.

In every case, the chi-square test showed a significant difference ($P < .05$) between the expected values and the obtained values, suggesting that the levels of training were not randomly distributed for each procedure. Table I shows the percentage of respondents who indicated that each procedure could be masterable at each level of training. Only procedures indicated for a particular level of training by 50% or more respondents are listed, except for PGY1 and PGY5 levels, where the most commonly cited procedures are listed. The second survey was composed of the most often-cited procedures (note that this table includes surveys returned after the second survey was for-

TABLE I.
Percentage of Respondents Indicating that Competency in Each Procedure could be Attained at a Particular Level of Training.

	PGY1	PGY2	PGY3	PGY4	PGY5
Adenoidectomy		64.5			
Anterior ribbon gauze pack	41.3	55.6			
Blepharoplasty, bilateral				50.8	
Bronchoscopy, diagnostic		52.4			
Cochlear implantation					41.3
Debridement of ears	42.9	50.8			
Esophagoscopy		65.1			
Laryngectomy, total				57.1	
Lip excision/flap reconstruction				52.4	
Lip-wedge resection			57.1		
Mandible fracture, closed reduction			49.2		
Mandible resection				50.8	
Mastoidectomy, MR				49.2	
Maxillectomy, medial				57.1	
Maxillectomy, total					55.6
Myringotomy and tube insertion		68.3			
Neck dissection, modified radical			55.6		
Neck dissection, selective				51.6	
Parathyroidectomy					45.2
Parotidectomy, superficial				48.4	
Peritonsillar abscess drainage		75.8			
Pharyngeal diverticulectomy				56.5	
Posterior nasal pack		69.4			
Rhinoplasty					53.2
Skin graft-split thickness	27.0				
Skin lesion excision		59.7			
Stapedectomy					59.7
Submandibular gland excision			54.8		
Thyroglossal duct cyst excision			66.1		
Thyroid lobectomy				62.9	
Thyroidectomy, total				53.2	
Tonsillectomy		74.2			
Tracheoesophageal puncture			56.5		
Tracheotomy		59.7			
Transnasal maxillary sinusotomy			50.0		
Turbinectomy		53.2			

POY = postgraduate year; MR = modified radical.

mulated so that there are minor differences in the rankings; for example, for PGY3, closed reduction of mandible fracture ultimately ranked just below transnasal maxillary sinusotomy and wedge resection of the lip, both of which were not included on the second survey).

Estimated number needed for competence ranged from 4 for partial auriculectomy to 15 for rhinoplasty. Table II shows estimated numbers and actual graduate case numbers from the ABO documents for 2002 for the 28 procedures that were included on the second survey.

Survey 2

The response rate for survey 2 was 69%, and respondents were not identical to the first group, so that opinions from 81 of the 102 program directors were submitted for at least one of the surveys. Again, chi-square tests showed a significant difference for every procedure ($P < .05$), indicating the nonrandom distribution of degree of agreement for each combination of PGY-level and procedure. Table II shows the list of procedures on the survey and the training levels where at least 49% of respondents indicated that they "totally agree(d)" that this procedure should be performed competently by trainees at this level. In 21 of the 28 procedures, the mode was "totally agree." In 12 of these, the mode represented less than 49% of the respondents, leaving 16 representative procedures.

DISCUSSION

In this modified Delphi technique, otolaryngology residency program directors were chosen as the expert group because they would have experience with high-stakes evaluation of residents' clinical work (specifically,

TABLE II.
Second Survey Procedures: Percentage of Respondents Who Agreed that Competency could be Attained in Each Procedure at a Particular Level of Residency, Estimated Number to Competence, and Mean Number Reported by ABO Graduates in 2002.

	Training Level*	Percent Who "Totally Agree" with This Level	Estimated Number to Competence	Mean Number Reported by 2002 ABO Graduates
Anterior ribbon gauze pack	PGY1	70	5.1	—
Debridement of ears	PGY1	59	7.4	—
Skin graft-split thickness	PGY1	57	5.0	15.9
Adenoidectomy	PGY2	84	10.4	114.1
Esophagoscopy	—	46 (PGY2)	8.4	52.2
Myringotomy and tube insertion	PGY2	83	10.1	175.0
Peritonsillar abscess drainage	PGY2	87	4.9	18.0
Posterior nasal pack	PGY2	79	4.5	—
Tonsillectomy	PGY2	84	9.4	127.1
Tracheotomy	PGY2	61	9.4	73.7
Mandible fracture, closed reduction	PGY3	64	6.3	7.2
Mastoidectomy	—	27 (PGY3)	Simple—8.6 modified radical—9.6	Simple—10.4 modified radical—6.6
Neck dissection, modified radical	—	14 (PGY3)	10.0	27.7
Submandibular gland excision	PGY3	54	5.5	10.1
Tracheoesophageal puncture	—	47 (PGY3)	5.7	6.8
Thyroglossal duct cyst excision	PGY3	57	5.0	5.4
Blepharoplasty, bilateral	—	23 (PGY4)	8.1	4.3
Laryngectomy-total	—	41 (PGY4)	7.2	8.0
Mandible resection	—	33 (PGY4)	6.7	9.2
Neck dissection, selective	—	46 (PGY4)	9.8	15.3
Parotidectomy, superficial	PGY4	49	8.7	16.3
Thyroid lobectomy	PGY4	59	9.2	16.3
Thyroidectomy, total	PGY4	49	8.8	11.6
Cochlear implantation	—	23 (PGY5)	7.4	4.6 (mixed code†)
Maxillectomy, total	—	47 (PGY5)	6.2	3.7
Parathyroidectomy	—	47 (PGY5)	9.3	5.4
Rhinoplasty	PGY5	53	15.2	16.5
Stapedectomy	—	40 (PGY5)	10.5	6.8

*Training Level where mastery is expected, demonstrated by a mode (including at least 49% of respondents to the survey) of "totally agree" with the indicated level.

†This ABO category includes middle and posterior fossa skull base surgery as well as cochlear implantation.

POY = postgraduate year; ABO = American Board of Otolaryngology.

deciding whether they were competent enough to graduate from training). The first survey used the existing list of procedures as recorded by the ABO but dropped all procedures where the average resident would perform one or fewer, in the expectation that these are, first of all, not the basic otolaryngology procedures, and second of all, not viable as benchmarks because of their infrequency in the curriculum. This left 73 procedures, still a long and unwieldy number for the survey. An approximate 50% listing of a particular level of training as appropriate for a particular procedure was chosen as a cutoff for the second survey, but in some cases, the procedures eventually included had higher or lower levels of agreement. Notably, no procedures were considered appropriate for PGY1 by 50% or more of respondents, and only three reached this level for PGY5. The former is explainable because in many programs, the PGY1 resident is in the general surgery stream, and it would be unreasonable to expect them to be able to acquire otolaryngologic skills. The latter is partially explainable because respondents expected most of the procedures to be mastered by the PGY4 level, leaving the least common procedures for consideration for the PGY5s. In most training programs, the simpler cases are attended by the junior residents, and the more complex ones by the more senior residents. This is the root of the results we see in this survey. Certain cases are widely seen as junior level cases and others as senior level cases.

On the second survey, the average number of each procedure performed by 2002 ABO graduates was listed, and in only one procedure in the PGY5 section was the actual number of cases reported by these graduates higher than the program directors' estimated number needed for competency. There were a few other examples of cases where the estimated number needed for competency was greater than the reported number (modified radical mastoidectomy, for example). As well, residents perform far in excess of the estimated number needed for tonsillectomy, adenoidectomy, and ventilation tube insertion. This might represent curricular time that could be directed to other areas, such as work on models for less common cases.

An estimation of cases required to be competent represents an opinion. The reluctance of this group to participate in answering this question may suggest their acknowledgment that residents learn at different rates and that the same procedure may be more difficult in particular patients. A resident who has performed 10 tonsillectomies would be expected to know the mechanics of the procedure but will not have optimal speed and will not have encountered the difficult tonsillectomy. Competence should not be confused with expertise, which comes from reflection on significant experience and the ability to adapt behavior to achieve the desired outcome.[7]

There were very few cases felt to be beyond resident level. The most commonly cited was cochlear implantation, in which 39.7% of respondents felt it was beyond the abilities of a resident. Cricoid split at 15.9%, stapedectomy at 8.1%, and bilateral blepharoplasty at 7.9% were the next most frequently cited.

The Accreditation Council for Graduate Medical Education Outcomes project focuses on authentic assessment of core competencies in medical training. Surgical skill is only one of many competencies our residents need to achieve. The Assessment section of the Outcomes project lists some of the aspects of authentic assessment: it must be valid, reliable, fair, feasible, and provide valuable information about a selected sample of the program objectives, and it must use multiple observations of behavior by multiple observers, using a predetermined minimum standard of success and ideally include multiple evaluation tools.[8] This study has helped to select a sample from this area of resident competency. It can be used as the basis of a multilevel residency surgical skills curriculum. The next step is to develop a useable teaching and evaluation toolkit surgical faculty can use either in the operating room or with affordable bench models to instruct and provide formative and summative evaluation to residents. In the mean time, residents can use these benchmark procedures to monitor their own surgical skills.

CONCLUSION

This two-step survey technique yielded 16 possible surgical procedures for which mastery is felt to be achievable at the various levels of training. Most procedures fell into the PGY2 to PGY4 levels. There is a possibility that the most complex procedures felt masterable at the PGY5 level are on average performed relatively infrequently.

Acknowledgments

The author thanks the program directors who participated in this survey, and to Beth Shultz, Kimberly Kent, and Donna Chupick, who assisted in the preparation and mailing of the surveys.

BIBLIOGRAPHY

1. Sidhu RS, Grober ED, Musselman LJ, Reznick RK. Assessing competency in surgery: where to begin? *Surgery* 2004;135: 6–20.
2. Grober ED, Hamstra SJ, Wanzel KR, et al. The educational impact of bench model fidelity on the acquisition of technical skill. *Ann Surg* 2004;240:374–381.
3. Hilmi OJ, White PS, McGurty DW, Oluwole M. Bronchoscopy training: is simulated surgery effective? *Clin Otolaryngol* 2002;27:267–269.
4. Mason TP, Applebaum EL, Rasmussen M, et al. Virtual temporal bone: creation and application of a new computer-based teaching tool. *Otol Head Neck Surg* 2000;122: 168–173.
5. Ward M, MacRae H, Schlacta C, et al. Resident self-assessment of operative performance. *Am J Surgery* 2003; 185:521–524.
6. Dillman DA. *Mail and Internet Surveys: The Tailored Design Method*, 2nd ed. Hoboken, NJ: John Wiley & Sons, 1999.
7. Guest CB, Regehr G, Tiberius RG. The life long challenge of expertise. *Med Educ* 2001;35:78–81.
8. Lynch DC, Swing SR. Key considerations for selecting assessment instruments and implementing assessment systems. ACGME Outcomes Project, on ACGME web site, 2001. Available at: http://www.acgme.org/outcome/assess/keyConsider.asp. Accessed January 4, 2005.