B-ENT, 2012, **8**, 191-195

Development of a Tool for Global Rating of Endoscopic Surgical Skills (GRESS) for Assessment of Otolaryngology Residents

O. Marglani[1], A. Alherabi[1], T. Al-Andejani[2], A. Javer[3], A. Al-Zalabani[4] and A. Chalmers[5]

[1]Department of Otolaryngology, Head & Neck Surgery, College of Medicine, Umm-al-Qura University, Makkah, Saudi Arabia; [2]Department of Otolaryngology, King Abdulaziz Medical City, Jeddah, Saudi Arabia; [3]Department of Otolaryngology, St. Paul's Sinus Centre, St. Paul's Hospital, Vancouver, British Columbia, Canada; [4]Department of Community Medicine, Taiba University, Madinah City, Saudi Arabia; [5]Department of Rheumatology and Medical Education, University of British Columbia, Vancouver, British Columbia, Canada

Key-words. Otolaryngology; endoscopic sinus surgery; tool; assessment; skills

Abstract. *Development of a Tool for Global Rating of Endoscopic Surgical Skills (GRESS) for Assessment of Otolaryngology Residents*. *Objective*: To develop a valid and reliable assessment tool for endoscopic sinus surgery (ESS).
Material and methods: Data were collected prospectively in an observational study through evaluations at two tertiary academic institutions, i.e. St. Paul's Sinus Centre, St. Paul's Hospital, Vancouver, British Columbia, Canada, and King Fahd Medical City, Riyadh, Saudi Arabia, from December 2006 to December 2009. A 2-page evaluation form was developed in conjunction with the Objective Assessment of Technical Skills Surgery (OSATS) evaluation form developed by Reznick et al in Toronto to assess residents' surgical skills. A Likert scale (1-5 where 5 = excellent) was used for evaluations. The Global Rating of Endoscopic Surgical Skills (GRESS) evaluation instrument was designed with input from academic otolaryngologists, fellowship-trained rhinologists, and experts in medical education. The experts' comments were incorporated, establishing face and content validity. Residents from various levels of training were assessed objectively using this instrument. Internal consistency was evaluated using Cronbach's α. Test-retest and inter-rater reliability was measured using intra-class correlation.
Results: A total of 31 assessments were completed by 15 residents. GRESS showed high reliability in the context of internal consistency ($\alpha = 0.99$), test-retest (0.95, CI = 0.83-0.98), and inter-rater reliability (0.86, CI = 0.31-0.98).
Conclusions: This pilot study demonstrated that GRESS is a valid and reliable assessment tool for operating room performance.

Introduction

Endoscopic sinus surgery (ESS) is an essential component of otolaryngology surgical training.[1] Endoscopic surgery is a reliable choice for removing even large, posteriorly-located, intranasal lesions, e.g. intranasal pleomorphic adenoma can be approached using this technique, without open surgery.[2] Competence has become an important issue in current surgical practice and training.[1] There is also a pressing need for an intraoperative assessment tool that meets the high standards of reliability and validity as an outcome measure for different training strategies, and for recruiting foreign graduates into countries.[3]

The goal of our study was to determine the validity and reliability of an assessment tool for ESS.

Materials and methods

Data were collected prospectively in an observational study through evaluations at two tertiary academic institutions, i.e. St. Paul's Sinus Centre, St. Paul's Hospital, Vancouver, British Columbia, Canada from December 2006 to December 2007 and King Fahd Medical City, Riyadh, Saudi Arabia, from January 2008 to December 2009. The study subjects were residents of the Canadian and Saudi board programs in otolaryngology.

A 2-page evaluation form was developed in conjunction with the Objective Assessment of Technical Skills (OSATS) for Surgery developed by Martin *et al*.[4] (Table 1) in Toronto to assess the surgical skills of residents. A Likert scale (1-5 where 5 = excellent) was used for the evaluation.

The Global Rating of Endoscopic Surgical Skills (GRESS) evaluation instrument (Table 2) was designed with input from academic otolaryngologists, fellowship-trained rhinologists, and experts in medical education. The experts'

Table 1
Tool 1; Likert Scale for Global Rating of Surgical Skills (Martin *et al.*[4])

	Rating				
	1	**2**	**3**	**4**	**5**
Criterion					
Respect for tissue	Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments	Careful handling of tissue, but occasionally caused inadvertent damage			Consistently handled tissues appropriately with minimal damage
Time & Motion	Many unnecessary moves	Efficient time/motion, but some unnecessary moves			Economy of movement and maximum efficiency
Instrument handling	Repeatedly makes tentative or awkward moves with instruments	Competent use of instruments, but occasionally appeared stiff or awkward			Fluid moves with instruments and no awkward movements
Knowledge of instruments	Frequently asked for the wrong instrument or used an inappropriate instrument	Knew the names of most instruments and used appropriate instrument for the task			Obviously familiar with the instruments required and their names
Flow of operation and forward planning	Frequently stopped operating or needed to discuss the next move	Demonstrated ability for forward planning with steady progression of operative procedure			Obviously-planned course of operation with effortless flow from one move to the next
Knowledge of specific procedure	Deficient knowledge. Needed specific instructions at most operative steps	Knew all important aspects of the operation			Demonstrated familiarity with all aspects of the operation
Use of assistants	Consistently placed assistants poorly or failed to use assistants	Good use of assistants most of the time			Strategically used assistants to the best advantage at all times

comments were incorporated, establishing face and content validity. Residents from various levels of training were assessed objectively using this instrument intra-operatively, and were provided constructive feedback. Internal consistency was evaluated using Cronbach's α. Test-retest and inter-rater reliability was measured using intra-class correlations. Data were analyzed using SPSS version 16 (SPSS Inc., Chicago, IL, USA).

Ethical Issues

The Institutional Review Boards of the institutions involved granted us permission to conduct this study. We declare that we have no financial or personal relationship(s) which may have inappropriately influenced us in writing this paper.

Results

A total of 31 assessments were completed for 15 residents who were evaluated by 5 faculty members as they performed ESS on patients over a period of 3 years from 2006-2009. Three residents were observed by 2 faculty members simultaneously.

Internal consistency, evaluated using Cronbach's α, showed high reliability (α = 0.99). Test-retest reliability measured using intra-class correlation coefficient was also found to be high (ICC = 0.95, CI = 0.83-0.98). The intra-class correlation coefficient for the inter-rater reliability was ICC = 0.86 (CI = 0.31-0.98). In this study, both instruments showed construct validity, with an overall trend toward a higher score with a more advanced postgraduate year of training, and the faculty substantially outperformed residents (Figure 1).

Discussion

It seems obvious that direct observation of surgical skills in the operating room represents the 'gold standard' in terms of content and construct validity, knowing that surgical knowledge is probably best assessed by examinations, but surgical skills are best assessed in the workplace. Log-books form a useful record of experience gained, but experience does not necessarily equate with competence; therefore, developing such tools is vitally important. A good assess-

Table 2
Tool 2; Likert Scale for Global Rating of Endoscopic Surgical Skills (GRESS)

Tool 2: Global Rating of Endoscopic Surgical Skills (GRESS)					
	Rating				
	Unsatisfactory	**Adequate/satisfactory**			**Excellent/competent**
	1	**2**	**3**	**4**	**5**
A Patient Setup					
1. Position of ETT	Not positioned properly.	Had to re-adjust during the procedure			Positioned well, no need for re-adjustments
2. Eye protection: Lube in the eyes. Uncovered	Unaware of eye protection. Didn't consider it	Improper coverage of the eyes.			Lube in the eyes and taped properly
3. Head/patient positioning	Unaware of proper position. Didn't consider it.	Improperly positioned, made some effort.			Appropriately positioned, appropriate rest, e.g egg foam crate.
4. Patient preparation: -Nasal decongestants -Throat pack	Unaware of patient preparation. Not considered.	Makes some effort to prepare patient. Placed improperly.			Ensures correct preparation of the patient: nasal decongestion, placement of throat pack
B Equipment Setup					
1. Proper setup of IGS/TV tower/endoscopes	Unaware of proper setup	Setup improperly done.			Setup properly done Endoscope: correct image size, focusing of image, white balance.
2. Proper draping and placement of IGS tower	Unaware, did not attempt	Improperly placed and draped			Properly placed and draped
C Use of endoscope					
1. Operative field	Operative field rarely central	Operative field usually central			Operative field always central
2. Perception of depth	Lack of perception of depth, causing trauma to structures	Incomplete perception of depth.			Complete perception of depth
3. Instrument visualization endoscopically	Rarely keeps tip of instrument in view.	Sometimes keeps instrument tip in view, but not consistently			Consistently keeps instrument in view
4. Communication with anesthesiologist when injecting local, moving head, BP control	No communication	Sometimes, not consistent			Consistent communication

ETT: Endotracheal tube; IGS/TV: Image guided surgery/television; IGS: image guided surgery.

ment must possess reliability, validity, educational impact, acceptability, and feasibility.[5]

Assessment can be defined as making a judgment against a defined reference, and it has two main purposes, which ideally should not interfere with each other. The first is to provide feedback to aid learning, i.e. a formative or low-stakes assessment, and the second is for examination/certification, i.e. a summative or high-stakes assessment. Competence-based assessments measure what a surgeon can do in a controlled representation of professional practice, e.g. when observed in the operating room or in a clinical skills laboratory.[6] Miller's triangle defines a simple hierarchy for the development and assessment of clinical skills. The GRESS is based on this model, and was developed by the modified Delphi technique, which uses email to gather information, provide feedback, and report conclusions.[7]

Simulations have been suggested to form the basis for technical skills training and assessment in the future. This is because of the decreasing opportunity to practice on real patients and the need for 'deliberate practice' in a non-threatening environment.[8] Practice and assessment on simulations are no substitute for operative experience, but they enable surgeons to

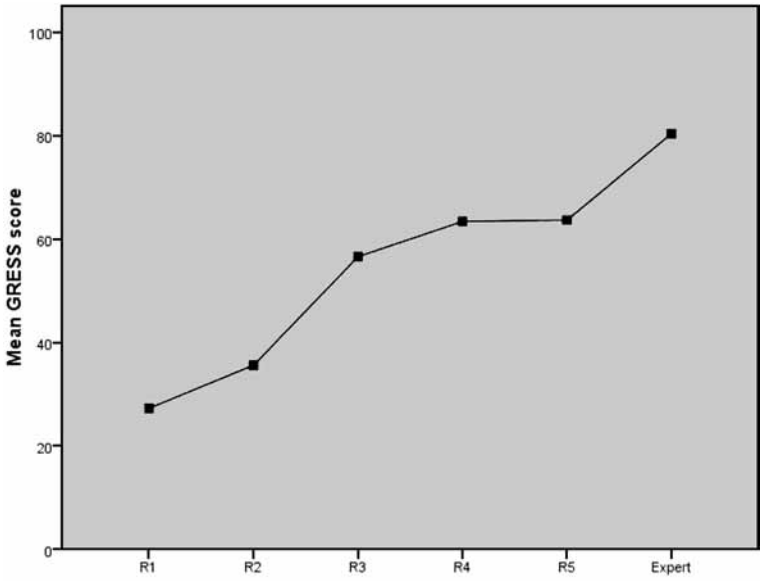

Figure 1
Assessment of the construct validity of the GRESS tool

become competent (and, therefore, confident) in key surgical skills before entering the complex operating room environment. A randomized trial showed that trainees who receive simulation training perform significantly better in the operating room.[9] Others have developed tools, but only at the cadaveric level.[10]

The faculty found the instrument to be easy to understand, complete, and practical and the residents felt that the instruments were helpful in providing immediate informative feedback on their performance.

In this study, we could not present data regarding complications that occurred with each resident during the evaluation while performing surgery. Actually, the GRESS assessment tool was not designed to measure the possible surgical complications that could occur with any senior surgeon, and every surgical procedure has its own risk of complications. What really matters is the rate of surgical complications, which could only be valid for clinical effectiveness measurements if calculated over a certain time period by repeating the same procedure a certain number of times. Because the assessment rate for each resident was limited, it not valid to take their complication rates into account to confirm the validity of testing.

Conclusions

This pilot study demonstrated GRESS to be a valid and reliable assessment tool of operating room performance to provide systematic and comprehensive feedback as part of the learning cycle. This tool was developed as an intraoperative instrument as opposed to other tools which used cadavers. Potential applications will be in tracking resident development throughout postgraduate training, and offering a structured means of certifying operative skills.

Limitations

Recommendations

Because of the limited number of subjects in this study, more data should be gathered by using our instrument on a larger scale, e.g. in other residency programs and over a longer period of time. Other assessment tools and checklists specific for otolaryngology are currently being developed for research and evaluation.

Acknowledgement

The authors are very thankful to Dr. Mian Usman Farooq, Senior Specialist, King Abdullah Medical City, Makkah, Saudi Arabia for his technical support.

References

1. Syme-Grant J, White PS, McAleer JP. Measuring competence in endoscopic sinus surgery. *Surgeon*. 2008;6(1):37-44.
2. Wu F, Huang CC, Fu CH, Chen YL, Lee TJ. Transnasal endoscopic surgery for intranasal pleomorphic adenomas. *B-ENT*. 2010;6(1):43-47.
3. Brown DJ, Thompson RE, Bhatti NI. Assessment of operative competency in otolaryngology residency: Survey of US Program Directors. *Laryngoscope*. 2008;118(10):1761-1764.
4. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84(2):273-278.
5. Darzi A, Smith S, Taffinder N. Assessing operative skill. Needs to become more objective. *BMJ*. 1999; 318(7188):887-888.
6. Schuwirth LW. Assessing medical competence: finding the right answers. *Clin Teach*. 2004;1(1):14-18.
7. Miller GE. The assessment of clinical skills, competence and performance.

Acad Med. 1990;65(9 Suppl):S63-S67.
8. Issenberg SB, Gordon MS, Gordon DL, Safford RE, Hart IR. Simulation and new learning technologies. *Med Teach*. 2001;23(1):16-23.
9. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomised clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg*. 2004;91(2):146-150.
10. Lin SY, Laeeq K, Ishii M, Kim J, Lane AP, Reh D, Bhatti NI. Development and pilot-testing of a feasible, reliable, and valid operative competency assessment tool for endoscopic sinus surgery. *Am J Rhinol Allergy*. 2009;23(3):354-359.

Osama Marglani, Assistant Professor
Department of Otolaryngology
Head & Neck Surgery,
College of Medicine,
Umm-al-Qura University,
Consultant Otolaryngology
King Abdullah Medical City, Makkah, Saudi Arabia
Tel.: 00966-566500350
Fax: 00966-25664663
E-mail: osamamergalani@yahoo.com