

# Toward an objective assessment of technical skills: a national survey of surgical program directors in Saudi Arabia

Abdullah Alkhayal<sup>1</sup>  
 Shahla Aldhukair<sup>2</sup>  
 Nahar Alselaime<sup>1</sup>  
 Salah Aldekhayel<sup>1</sup>  
 Sultan Alhabdan<sup>1</sup>  
 Waleed Altaweel<sup>3</sup>  
 Mohi Elden Magzoub<sup>4</sup>  
 Mohammed Zamakhshary<sup>1,2</sup>

<sup>1</sup>Department of Surgery, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia;

<sup>2</sup>Public Health Section, King Abdullah International Medical Research Center, Riyadh, Saudi Arabia; <sup>3</sup>Urology Department, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; <sup>4</sup>Department of Medical Education, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia

**Background:** After almost a decade of implementing competency-based programs in post-graduate training programs, the assessment of technical skills remains more subjective than objective. National data on the assessment of technical skills during surgical training are lacking. We conducted this study to document the assessment tools for technical skills currently used in different surgical specialties, their relationship with remediation, the recommended tools from the program directors' perspective, and program directors' attitudes toward the available objective tools to assess technical skills.

**Methods:** This study was a cross-sectional survey of surgical program directors (PDs). The survey was initially developed using a focus group and was then sent to 116 PDs. The survey contains demographic information about the program, the objective assessment tools used, and the reason for not using assessment tools. The last section discusses the recommended tools to be used from the PDs' perspective and the PDs' attitude and motivation to apply these tools in each program. The associations between the responses to the assessment questions and remediation were statistically evaluated.

**Results:** Seventy-one (61%) participants responded. Of the respondents, 59% mentioned using only nonstandardized, subjective, direct observation for technical skills assessment. Sixty percent use only summative evaluation, whereas 15% perform only formative evaluations of their residents, and the remaining 22% conduct both summative and formative evaluations of their residents' technical skills. Operative portfolios are kept by 53% of programs. The percentage of programs with mechanisms for remediation is 29% (19 of 65).

**Conclusion:** The survey showed that surgical training programs use different tools to assess surgical skills competency. Having a clear remediation mechanism was highly associated with reporting remediation, which reflects the capability to detect struggling residents. Surgical training leadership should invest more in standardizing the assessment of surgical skills.

**Keywords:** objective assessment, program directors, surgical skills, surgical residency

## Background

There has been a recent, significant shift in residency training toward competency-based programs, such as the Canadian Medical Education Directives for Specialists (CanMEDS) framework<sup>1</sup> and the Accreditation Council for Graduate Medical Education (ACGME) Outcome Project.<sup>2</sup> These projects have identified certain competencies that residents should acquire during their training (Table 1). Surgical competency is covered under the medical expert and patient care sections in the CanMEDS and ACGME frameworks, respectively. However, after the initiation of these projects, program directors (PDs) have been struggling to identify an optimal assessment tool

Correspondence: Mohammed Zamakhshary  
 Associate Professor of Pediatric Surgery  
 Public Health Section, King Abdullah  
 International Medical Research Center,  
 King Saud bin Abdulaziz University for  
 Health Sciences, Riyadh, Saudi Arabia  
 Tel +966 543 427 669  
 Email zamakhshary@gmail.com

**Table 1** CanMEDS and ACGME core competencies

CanMEDS	ACGME
<ul style="list-style-type: none"> <li>• Medical expert</li> <li>• Communicator</li> <li>• Collaborator</li> <li>• Manager</li> <li>• Health advocate</li> <li>• Scholar</li> <li>• Professionalism</li> </ul>	<ul style="list-style-type: none"> <li>• Medical knowledge</li> <li>• Communication and interpersonal skills</li> <li>• Patient care</li> <li>• System-based practice</li> <li>• Practice-based learning and improvement</li> <li>• Professionalism</li> </ul>

**Abbreviations:** CanMEDS, Canadian Medical Education Directives for Specialists; ACGME, Accreditation Council for Graduate Medical Education.

to measure this competency that meets certain criteria.<sup>3,4</sup> The first criterion, reliability, refers to the reproducibility, consistency, and generality of the results over time. The second criterion, feasibility, reasons that, in order to be accepted, a new tool should be cost-effective, easy to execute, and not time-consuming. The third criterion, validity, asks if the tool assesses what it is supposed to assess. Validity is divided into five sub-categories: predictive validity, which asks if the tool can predict future performance; face validity, which asks if the tool reflects real life; construct validity, which asks if the tool measures what it is designed to measure; content validity, which asks if the domain that we are assessing is assessed by the tool; and criterion-related validity, which asks if the result assessed by a tool correlates with that measured by the current gold standard tool.

After almost a decade of implementing competency-based programs in postgraduate training programs, the assessment of technical skills remains more subjective than objective, relying mainly on the gut feelings of the assessor and on nonstandardized direct observation of the resident while the resident performs a procedure. These subjective assessment methods have poor reliability, as manifested by measuring interrater and test–retest reliability.<sup>3,5</sup> Dr Carlos A Pellegrini, the former president of the American Surgical Association, questioned in his presidential letter: “Is performance measurement possible and accurate?”<sup>6</sup> He sought a valid, reliable, measurable, and feasible assessment tool. In recent decades, surgical educators around the world have developed and validated objective assessment tools to measure surgical residents’ technical skills. Martin et al invented a new method that resembles the Objective Structured Clinical Examination and named it the Objective Structured Assessment of Technical Skills (OSATS).<sup>7</sup> The future of training is moving toward the use of more technology; Fried et al<sup>8</sup> developed the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS). In addition, Datta et al<sup>9</sup> developed the Imperial College Surgical Assessment Device

(ICSAD), which assesses surgical residents’ dexterity by electromagnetic means.

The Saudi Commission For Health Specialties (SCFHS)<sup>10</sup> was established under a royal decree in 1983. Its responsibilities are to organize, evaluate, and set policies for training program accreditation; in addition, the SCFHS is the accreditation and qualifying agency of the trainees. Thirty-seven residency and fellowships training programs in multiple health specialties have been recognized by the SCFHS. One of the main aims of the SCFHS is to supervise and assure the quality of graduates by offering a fair assessment to all trainees.

To enact a new change in a complex organization system, Heifetz and Laurie<sup>11</sup> recommend “getting on the balcony” to have a clear view of the current situation. Data on the assessment of technical skills in surgical specialties in national settings are lacking. We conducted this research to explore the current tools used by different surgical programs for technical skills assessment, their relationship with remediation (ie, the action taken by a program director toward a resident who does not fulfill the criteria for promotion to the next level of training), the recommended tools from the PDs’ perspective, and PDs’ attitudes toward the available objective tools for technical skills assessment.

## Methods

### Study design

A cross-sectional study was carried out between January 27, 2010 and March 24, 2010 and involved all of the surgical PDs accredited by the SCFHS.

### Population, sample size, and sampling

This census study involved all of the PDs of all surgical programs accredited by the SCFHS (general surgery, orthopedics surgery, urology, otolaryngology, plastic surgery, maxillofacial surgery, obstetrics, and gynecology). The inclusion criterion was a surgical PD accredited by the SCFHS. The exclusion criteria were PDs not identified by the SCFHS or with invalid or incorrect email addresses on the SCFHS website. Ophthalmology and fellowship PDs were also excluded.

### The survey

For the survey, we used a modified version of a previously used survey after receiving author permission.<sup>12</sup> The modified survey was designed by the primary investigator and the coauthors. A focus group was developed to assess the content and face validity of the survey. The focus group was

composed of five surgeons from different surgical specialties who hold degrees in epidemiology and/or medical/surgical education. These surgeons were non-PDs to avoid contamination of the final sample.

The survey contains demographic information about the program (specialty, number of residents evaluated each year, number of residents accepted each year, number of full-time staffers, and the presence of a simulation lab). Respondents were provided a list of the available objective assessment tools with a brief description of each and asked to identify which of the tools they used and explain the reason for not using any of the available tools, if applicable. Remediation was also covered (the presence of a clear remediation mechanism based on poor surgical performance, information on when it takes place, and the number of residents that required remediation). The last section involves the importance and the recommended tools from the respondent's perspective, his/her attitudes and motivations to apply those tools in his/her training program, and any further comments.

The survey was sent to the participants' emails through the SurveyMonkey® website. A reminder email was sent to participants who failed to respond within 2 weeks.

## Statistical analysis

All data were collected electronically and downloaded as an Excel (Microsoft Corporation, Redmond, WA) spreadsheet. From there, they were incorporated into the SPSS program (v 17; IBM Corporation, Armonk, NY). Descriptive statistics were tabulated, as were mean and median calculations for the scoring of various questions. The response to each question was considered as a continuous variable, and Chi-squared and Fisher's exact tests were used where appropriate.

We then conducted a correlation analysis using the Pearson correlation coefficient with the remediation rate for each assessment tool. The associations were also calculated using odds ratios for each assessment tool. Both *P*-values less than 0.05 and 95% confidence intervals that did not overlap with 1 were considered statistically significant.

## Results

### Response rate

A total of 116 PDs were surveyed in the study, and 71 directors responded to our survey, for a response rate of 61%. Of these 71 directors, 65 respondents (56%) completed the survey. Six incomplete responses were excluded. The response rate after the first mailing was 18 (16% of the total) respondents; after the second mailing, another 53 responded (46%).

## Demographic data

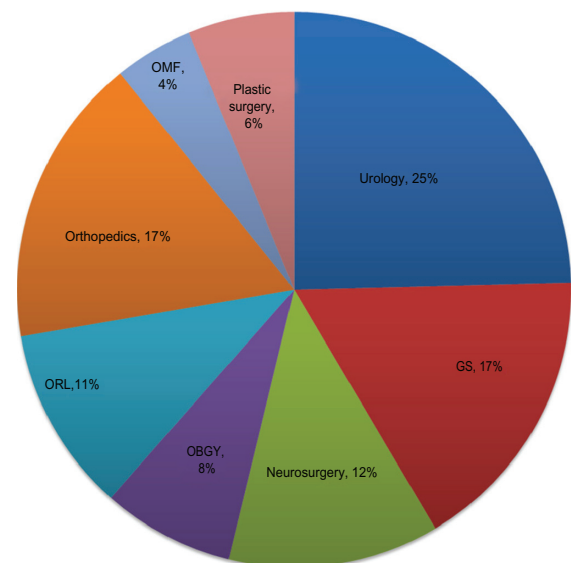
Urology PDs comprised 25% ( $n = 16$ ) of the respondents (Figure 1). The mean number of residents evaluated each year by a single PD was 11.4 (median = 8), with the highest number in general surgery. The mean number of residents accepted each year per institution was 2.74 (median = 2), with the highest number in general surgery. Fourteen percent of PDs reported having a simulation lab in their institution. The complete descriptive characteristics of the studied programs are shown in Table 2.

## Surgical skills evaluation and feedback

Of the respondents, 39 (60%) only performed summative evaluations of their residents (ie, end of rotation/year evaluation), whereas 10 (15%) only conducted formative evaluations of their residents (ie, continued evaluation throughout the rotation/year). The remaining 14 (22%) used both summative and formative evaluations of residents' technical skills. The evaluation and feedback was provided orally by 37 (57%) respondents and was written for 48 (74%) respondents. The most common evaluation used was an end-of-rotation, written summative evaluation (70%). The distribution of evaluation techniques by specialty is shown in Figure 2.

## Technical skills assessment tools

Of the respondents, 42% thought that they were using more than subjective technical skills assessment, whereas the remaining 59% mentioned using only nonstandardized,



**Figure 1** Distribution of respondents by specialty.

**Abbreviations:** GS, general surgery; OBGY, obstetrics and gynecology; ORL, otorhinolaryngology; OMF, oral maxillofacial surgery.

**Table 2** Demographic data of the programs surveyed

Specialty (number of respondents/number of total program directors surveyed)	Mean number of residents evaluate/year (range)	Mean number of residents accepted each year (range)	Mean number of consultants (range)	Consultant to resident ratio	Number of programs with a simulation lab
Urology (n = 16/20)	5.6 (1–20)	1.18 (1–2)	5.56 (2–11)	1:1	4
Neurosurgery (n = 8/11)	5.5 (1–8)	2 (1–4)	4.5 (3–12)	1:1.2	2
GS (n = 11/18)	26 (5–60)	3.9 (2–6)	11.63 (5–24)	1: 2.24	3
OBGY (n = 5/16)	15.2 (4–22)	3.2 (2–5)	13.4 (7–21)	1:1.33	2
OMF (n = 3/8)	7.33 (4–12)	3 (2–5)	4 (3–5)	1:1.83	0
ORL (n = 7/14)	8.1 (3–12)	3.71 (1–10)	5.86 (3–14)	1:1.38	1
Plastic surgery (n = 4/5)	10.75 (2–20)	3.5 (1–6)	4.25 (3–6)	1: 2.53	1
Orthopedics (n = 11/24)	14.36 (4–30)	3.18 (1–10)	7.3 (3–12)	1:1.97	1

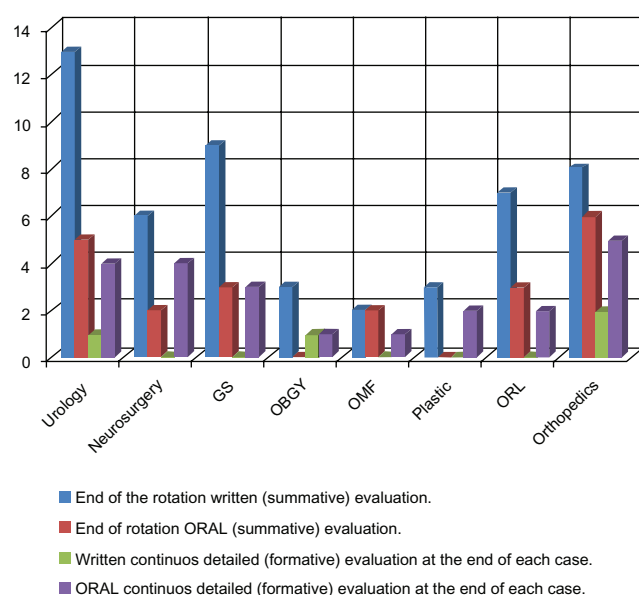
**Abbreviations:** GS, general surgery; OBGY, obstetrics and gynecology; ORL, otorhinolaryngology; OMF, oral maxillofacial surgery.

subjective, direct observation for technical skills assessment. Eighteen (26%) respondents were using at least one of the assessment tools listed in the survey (ie, OSATS, videotaped analysis, and resident surgical portfolios). Resident surgical portfolios refer to the log books containing all surgical cases in which the resident participated, either as the primary surgeon or the assistant, during his/her residency training. Eight (12%) respondents reported using hand-motion analysis, and five (8%) respondents reported using simulation to assess

residents' technical skills. Table 3 shows the distribution of assessment tools used by specialty.

## OSATS

Seven (11%) respondents were using the OSATS, and orthopedic surgeons used it the most. Those respondents who never used it (59, 89%) attributed their non-use to the fact that it was not validated in their specialty (23, 39%), that they were not aware of the supporting literature (21, 36%), and that it was either not cost-effective (3, 5%) or not practical (6, 10%). In addition, eight respondents (14%) attributed its non-use to availability issues, whereas four (7%) attributed its non-use to lack of institutional support and resources. Of the respondents, 80% judged the OSATS as important, whereas only 1.5% thought that it was not important. Furthermore,

**Figure 2** Distribution of evaluation technique by specialty.

**Abbreviations:** GS, general surgery; OBGY, obstetrics and gynecology; ORL, otorhinolaryngology; OMF, oral maxillofacial surgery.

**Table 3** Distribution of assessment tools used by specialty

Specialty	OSATS	Hand motion analysis	Videotaped analysis	Residents surgical portfolios
GS	1	2	4	5
Urology	1	2	3	10
ORL	1	0	1	3
Neurosurgery	0	1	0	3
OMF	0	1	1	3
Plastic	0	1	0	4
OBGY	1	1	0	2
Orthopedics	3	0	1	5

**Abbreviations:** OSATS, Objective Structured Assessment of Technical Skills; GS, general surgery; OBGY, obstetrics and gynecology; ORL, otorhinolaryngology; OMF, oral maxillofacial surgery.

5% and 14% stated that they did not know or were neutral regarding the importance, respectively. Many participants (85%) recommended that the OSATS be included as part of a technical skills assessment program, whereas the other participants were neutral (11%) or did not know (5%).

## Videotaped analysis

Ten (15%) respondents used videotaped analysis, and general surgery programs used this tool the most. Of those respondents who never used it (55, 85%), reasons given were that it was not validated in their specialty (21, 38%); they were not aware of the supporting literature (19, 35%); it was not cost-effective (7, 13%), practical (14, 26%), or available (4, 7%), and other reasons (8, 15%). Of the respondents, 60% thought that videotaped analysis was important; 32% thought it had no effect, whereas 6% thought that it was not important. Many of the participants (63%) recommended that videotaped analysis be part of a technical skills assessment program, whereas 29% were neutral, and 3% did not recommend its use.

## Resident surgical portfolios

A total of 35 (54%) respondents used resident surgical portfolios. Urology PDs used them the most. Of those respondents who never used these portfolios (30, 46%), the reasons provided were that they were not validated in their specialty (11, 37%), that they were not aware of the supporting literature (12, 40%), that they were either not cost-effective (3, 10%) or practical (5, 17%), and other reasons (5, 17%). Of the respondents, 79% thought that resident surgical portfolios were important and should be part of a technical skills assessment program, whereas only 19% thought that portfolios were neutral. No respondents thought that portfolios were not important or did not recommend their use.

## Remediation

Nineteen programs reported having a clear mechanism for remediation based on poor surgical performance. These remediation mechanisms take place at the end of the rotation, at the end of the academic year, and at the end of training at frequencies of 68%, 26%, and 5%, respectively. Thirty-one (48%) PDs reported having had at least one remediation case in the past 5 years based on poor surgical performance. The mean number of residents in all programs remediated in the past five years was 1.77 (1–4), and a total of 55 residents were required to remediate. Using videotaped analysis as an objective assessment tool was significantly associated with reporting remediation ( $P = 0.026$ ), but there was no

significant association with the use of the OSATS or surgical residents' portfolios and remediation (Table 4). However, having a clear remediation mechanism based on poor surgical performance was highly significantly associated with reporting remediation ( $P = 0.001$ ). However, no association was identified between reporting remediation and the different modes of evaluation and feedback.

## Satisfaction, attitude, motivation

Only 39% of respondents are satisfied (Figure 3) with the methods they have been using to assess technical skills, and this satisfaction was significantly associated with the use of surgical residents' portfolios ( $P = 0.035$ ). Of the respondents, 91% agreed that our survey increased their knowledge about the objective assessment of technical skills; 89% and 88% of the respondents agreed that our survey motivated them to learn more about and to apply objective assessment tools in their programs, respectively. A total of 94% of the participants expressed interest in attending a workshop on objective assessment tools. The comments provided by the participants were very favorable, and, in general, the participants found the survey to be extremely helpful. As an example of how participants valued the survey, one participant stated, "All members of examination committees in different programs should have continuous updates of their methods of evaluation." Most comments also involved expressing interest in attending any further workshops or training pertaining to the matter of the objective assessment of surgical skills, e.g. "I think we need a workshop on acceptable methods for all trainers in all programs in Saudi Arabia."

## Discussion

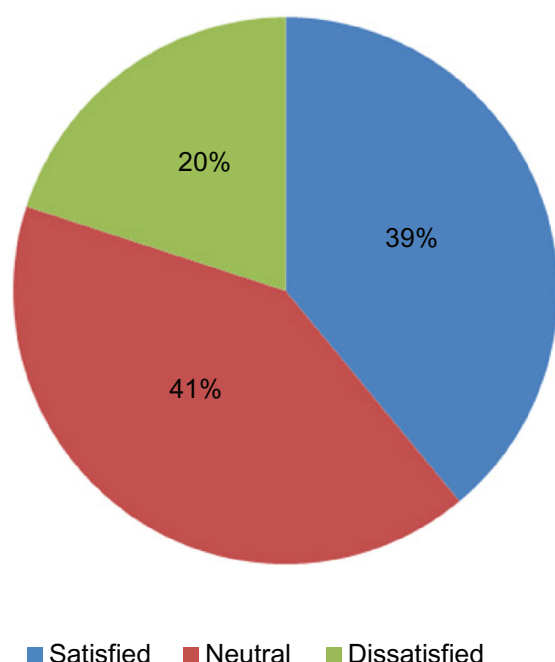
Technical surgical errors comprise the majority of surgical errors, with more than half occurring due to a lack of

**Table 4** The association between remediation and both different assessment tools and the presence of a remediation mechanism

		Remediation		Odds ratio (95% CI)	P-value
		Yes	No		
OSATS	Yes	4	3	1.53 (0.314–7.457)	0.596
	No	27	31		
Videotaped analysis	Yes	8	2	9.85 (1.080–28.675)	0.026
	No	23	32		
Surgical residents' portfolios	Yes	17	18	1.08 (0.406–2.865)	0.878
	No	14	16		
Remediation Mechanism	Yes	15	4	7.03 (1.997–24.759)	0.001
	No	16	30		

**Abbreviations:** CI, confidence interval; OSATS, Objective Structured Assessment of Technical Skills.





**Figure 3** Program directors' satisfaction regarding technical skills assessment.

competence.<sup>13,14</sup> Reports have suggested that objective assessments be used to minimize medical errors.<sup>15</sup> The SCFHS pays a great deal of attention to standardizing training and assures the quality control of graduates by adapting specific criteria for the assessment of the quality of training centers and by conducting annual national exams for all postgraduate medical training programs in the different specialties.<sup>10</sup> Despite this great interest in highly standardized training programs, national exams only assess clinical skills and knowledge by using multiple choice questions and an Objective Structured Clinical Examination, whereas the assessment of surgical residents' technical skills is completely ignored and left to each individual institution. Due to this nonstandardized technical skills assessment, graduate residents' surgical competency varies widely. We hope our study will serve as a cornerstone for developing a standardized objective assessment program for surgical residents' technical skills, both nationally and internationally.

The detection of struggling residents during training is one of the main aims of PDs. Subjective evaluation, which has poor reliability,<sup>3,5</sup> may not detect this issue. Therefore, the use of objective assessment tools such as the OSATS, MISTELS, and electromagnetic hand motion analysis could be a "safeguard" against this dilemma, as all trainees would be tested fairly and objectively using validated tools.<sup>5,16</sup> Applying these tools could be another struggle; most of our respondents are motivated and willing to apply new

and innovative methods for assessment but lack knowledge and information about these tools. Thus, adapting faculty enhancement programs directed toward surgical education could solve this problem.

Having clear policies and mechanisms for remediation is helpful in detecting struggling trainees. The association between implementing such policies and reporting residents requiring remediation is confirmed in this study. However, the use of objective assessment tools did not display any association with reporting remediation, possibly because they are relatively new tools and their implementation is recent. Videotaped analysis is not a consistent tool, but using it with a validated checklist increases its reliability, validity, and objectivity.<sup>17</sup> However, if used without checklists, subjective direct observation continues to have disadvantages.

Bandiera et al<sup>18</sup> preferred using simulations, direct observation and in-training evaluation reports, logbooks, and objective structured assessment tools such as the OSATS and MISTELS for technical skills assessment. The ACGME<sup>2</sup> recommend that checklists and 360° evaluation be the first methods used for performance assessments of medical procedures. However, most of our participants recommended that the OSATS and residents' surgical portfolios be used first. Moreover, although more than half of the PDs surveyed use portfolios, the quality of these portfolios is unknown. In addition, the literature has demonstrated poor reliability and validity for this instrument.<sup>3</sup> Therefore, the assessment should not depend primarily on portfolios, and other assessment tools should be included;<sup>19</sup> portfolios provide an idea of the quantity, but not the quality, of the procedures performed.<sup>3</sup> The reliability and construct validity of the OSATS are established,<sup>20</sup> but only a minority of programs use it, either due to a lack of knowledge and/or inadequate resources.

Recent reports on incorporating technology and health informatics into postgraduate training and assessment are fascinating. Reports from Tufts University<sup>21</sup> on adapting the OpRate application to follow the progress of surgical residents' technical skills demonstrate that adapting such formative written application programs is useful in identifying residents who need special consolidated training in one area or another. Another reliable application adapted by Southern Illinois University<sup>22</sup> is the operative performance rating system, which is an internet-based, procedure-based assessment including ten checklist items that are used to evaluate residents' technical and decision-making skills for six different general surgical procedures.

Dexterity-based analysis devices have been covered extensively in the literature. The ICSAD has been validated as a hand motion analyzer in both open and laparoscopic surgeries.<sup>23,24</sup> The ICSAD tracks hand motion efficiency by measuring the number and speed of hand motions, the time taken to complete the task, and the distance traveled by the surgeon's hand. The Advanced Dundee Endoscopic Psychomotor Tester<sup>25</sup> is a computer-controlled endoscopic performance assessment system that measures the time taken to complete the task, the contact error time, the route of the props, and the number of completed tasks within a time limit. This system has reasonable reliability<sup>26</sup> and contrast validity<sup>27</sup> as an endoscopic assessor. The third system, the Hiroshima University Endoscopic Surgical Assessment Device measures the time taken to complete the task and the deviation of the instrument tip in the horizontal and vertical axes. Its initial reliability and validity reports are encouraging.<sup>28</sup>

Simulators have gained great popularity in training and, recently, in assessment. The minimally invasive surgical trainer-virtual reality (MIST-VR) simulator is one of the first simulators developed for training basic laparoscopic skills.<sup>29</sup> MIST-VR has been validated for assessing basic laparoscopic skills.<sup>30</sup> Maithel et al<sup>31</sup> compared three commercially available virtual reality simulators, MIST-VR, Endotower, and the Computer Enhanced Laparoscopic Training System (CELTS). All three simulators met the criteria for construct and face validity, and CELTS had the highest scores. Fried et al<sup>8</sup> also developed MISTELS, which includes five laparoscopic task stations (peg transfer, precision pattern cutting, securing a ligating loop, intracorporeal knotting, and extracorporeal knotting) performed in standardized laparoscopic training boxes. The reliability and validity of this system have been confirmed.<sup>8,32–36</sup> In addition, MISTELS has been shown to be superior to subjective in-training evaluation reports in identifying residents with poor laparoscopic skills.<sup>37</sup> Recently, a cognitive component was incorporated into MISTELS through the fundamentals to laparoscopic surgery (FLS) module, and FLS has been reported to have reasonable reliability.<sup>38</sup> However, the validity of the FLS is questionable<sup>39</sup> because the cognitive element of the assessment cannot differentiate between experts and novices.

Few published reports have focused on the real situation of technical skills assessment. Brown et al<sup>12</sup> surveyed otolaryngology PDs in the United States and found that most of the participants used subjective evaluation. Similarly, they identified a significant correlation between having a clear remediation mechanism and the remediation rate, which confirms that implementing clear mechanisms and

policies will enhance the identification of residents requiring remediation.

This study has several limitations. First, as in any survey study, the response rate is a major limitation and affects the calculated response rate for each specialty. Therefore, we sought to overcome this issue by frequently sending the survey to the participants, communicating via phone and on-site visits, and extending the survey time from 4 weeks to 6 weeks. Second, we assumed that the majority of participants would have a poor understanding of all of the assessment tools, so we added a brief description of each assessment tool in the survey, which may have caused a biased response. Third, the variety of specialties surveyed and the limited number of available programs make it difficult to generalize the results. Fourth, although we developed this survey scientifically using focus group methods and validated it prior to distribution to our target population, the survey has not been formally validated. Using the survey in a different population, such as with North American surgical PDs or any other regional PDs, would add credence to our results and will further validate both the survey and the results of our study.

## Conclusion

The survey revealed that surgical training programs use different tools to assess surgical skills competency. Having a clear remediation mechanism was highly associated with reporting remediation, which reflects the capability to detect struggling residents. Surgical training leadership should invest more in standardizing the assessment of surgical skills.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. CanMEDs Framework [webpage on the Internet]. Ottawa: Royal College of Physicians and Surgeons of Canada; c2012 [cited 2012 Jun 3]. Available from: <http://www.royalcollege.ca/portal/page/portal/rc/can-meds/framework>. Accessed February 22, 2010.
2. Toolbox of Assessment Methods [webpage on the Internet]. Chicago: Accreditation Council for Graduate Medical Education; c2000–2012 [cited 2012 Jun 3]. Available from: <http://www.acgme.org/acgmeweb/>. Accessed March 15, 2010.
3. Reznick RK. Teaching and testing technical skills. *Am J Surg*. 1993; 165(3):358–361.
4. Shumway JM, Harden RM; for Association for Medical Education in Europe. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach*. 2003;25(6):569–584.
5. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ*. 2003;327(7422):1032–1037.
6. Pellegrini CA. Surgical education in the United States: navigating the white waters. *Ann Surg*. 2006;244(3):335–342.

7. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84(2):273–278.
8. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg*. 2004;240(3):518–525; discussion 525–518.
9. Datta V, Mackay S, Darzi A, Gillies D. Motion analysis in the assessment of surgical skill. *Comput Methods Biomech Biomed Engin*. 2001; 4(6):515–523.
10. The Saudi Commission for Health Specialities [webpage on the Internet]. Riyadh: The Saudi Commission for Health Specialities; c2012 [cited 2012 Jun 3]. Available from: [http://english.scfhs.org.sa/index.php?option=com\\_content&view=article&id=100&Itemid=4378](http://english.scfhs.org.sa/index.php?option=com_content&view=article&id=100&Itemid=4378). Accessed April 2, 2010.
11. Heifetz RA, Laurie DL. The Work of Leadership. *Harvard Business Review*. 2001.
12. Brown DJ, Thompson RE, Bhatti NI. Assessment of operative competency in otolaryngology residency: Survey of US Program Directors. *Laryngoscope*. 2008;118(10):1761–1764.
13. Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery*. 2003;133(6):614–621.
14. Rogers SO Jr, Gawande AA, Kwaan M, et al. Analysis of surgical errors in closed malpractice claims at 4 liability insurers. *Surgery*. 2006;140(1):25–33.
15. Kohn LT, Corrigan JM, Donaldson MS. *To Err is Human: Building a Safer Health System*. Washington: National Academy Press; 1999.
16. Jaffer A, Bednarz B, Challacombe B, Sriprasad S. The assessment of surgical competency in the UK. *Int J Surg*. 2009;7(1):12–15.
17. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg*. 2008;247(2):372–379.
18. Bandiera G, Sherbino J, Frank, JR. *The CanMEDS Assessment Tools Handbook. An Introductory Guide to Assessment Methods for CanMEDS Competencies*. 1st ed. Ottawa: The Royal College of Physicians and Surgeons of Canada; 2006.
19. Tochel C, Haig A, Hesketh A, et al. The effectiveness of portfolios for post-graduate assessment and education: BEME Guide No 12. *Med Teach*. 2009;31(4):299–318.
20. Winkel CP, Reznick RK, Cohen R, Taylor B. Reliability and construct validity of a structured technical skills assessment form. *Am J Surg*. 1994;167(4):423–427.
21. Wohaibi EM, Earle DB, Ansanitis FE, Wait RB, Fernandez G, Seymour NE. A new web-based operative skills assessment tool effectively tracks progression in surgical resident performance. *J Surg Educ*. 2007;64(6):333–341.
22. Larson JL, Williams RG, Ketchum J, Boehler ML, Dunnington GL. Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents. *Surgery*. 2005;138(4):640–647; discussion 647–649.
23. Datta V, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A. Relationship between skill and outcome in the laboratory-based model. *Surgery*. 2002;131(3):318–323.
24. Mackay S, Datta V, Mandalia M, Bassett P, Darzi A. Electromagnetic motion analysis in the assessment of surgical skill: Relationship between time and movement. *ANZ J Surg*. 2002;72(9):632–634.
25. Hanna GB, Drew T, Clinch P, Hunter B, Cuschieri A. Computer-controlled endoscopic performance assessment system. *Surg Endosc*. 1998;12(7):997–1000.
26. Francis NK, Hanna GB, Cuschieri A. Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks. *Arch Surg*. 2001;136(1):40–43.
27. Francis NK, Hanna GB, Cuschieri A. The performance of master surgeons on the Advanced Dundee Endoscopic Psychomotor Tester: contrast validity study. *Arch Surg*. 2002;137(7):841–844.
28. Egi H, Okajima M, Yoshimitsu M, et al. Objective assessment of endoscopic surgical skills by analyzing direction-dependent dexterity using the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD). *Surg Today*. 2008;38(8):705–710.
29. McCloy R, Stone R. Science, medicine, and the future. Virtual reality in surgery. *BMJ*. Oct 20 2001;323(7318):912–915.
30. Taffinder N, Sutton C, Fishwick RJ, McManus IC, Darzi A. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator. *Stud Health Technol Inform*. 1998;50:124–130.
31. Maithel S, Sierra R, Korndorffer J, et al. Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? *Surg Endosc*. 2006;20(1):104–112.
32. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. *Am J Surg*. 1998;175(6):482–487.
33. Ghitulescu GA, DA, Feldman LS, et al. A model for evaluation of laparoscopic skills: is there correlation to level of training? *Surg Endosc*. 2001;15(S127).
34. Derossis AM, Bothwell J, Sigman HH, Fried GM. The effect of practice on performance in a laparoscopic simulator. *Surg Endosc*. 1998;12(9):1117–1120.
35. Peters JH, Fried GM, Swanstrom LL, et al. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery*. 2004;135(1): 21–27.
36. Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc*. 2003;17(6):964–967.
37. Feldman LS, Hagarty SE, Ghitulescu G, Stanbridge D, Fried GM. Relationship between objective assessment of technical skills and subjective in-training evaluations in surgical residents. *J Am Coll Surg*. 2004;198(1):105–110.
38. Swanstrom LL, Fried GM, Hoffman KI, Soper NJ. Beta test results of a new system assessing competence in laparoscopic surgery. *J Am Coll Surg*. 2006;202(1):62–69.
39. Zheng B, Hur HC, Johnson S, Swanström LL. Validity of using Fundamentals of Laparoscopic Surgery (FLS) program to assess laparoscopic competence for gynecologists. *Surg Endosc*. 2010;24(1): 152–160.

## Advances in Medical Education and Practice

### Publish your work in this journal

Advances in Medical Education and Practice is an international, peer-reviewed, open access journal that aims to present and publish research on Medical Education covering medical, dental, nursing and allied healthcare professional education. The journal covers undergraduate education, postgraduate training and continuing medical education

Submit your manuscript here: <http://www.dovepress.com/advances-in-medical-education-and-practice-journal>

Dovepress

including emerging trends and innovative models linking education, research, and healthcare services. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.