

## ORIGINAL ARTICLE

# Endoscopic surgical skill qualification system in Japan: Five years of experience in the gastrointestinal field

T Kimura<sup>1,2</sup>, T Mori<sup>1,3</sup>, F Konishi<sup>1,4</sup> & M Kitajima<sup>1,5</sup><sup>1</sup> The Japan Society for Endoscopic Surgery, Tokyo, Japan<sup>2</sup> Department of Surgery, Fujinomiya City General Hospital, Fujinomiya, Japan<sup>3</sup> Department of Surgery, Kyorin University, School of Medicine, Mitaka, Japan<sup>4</sup> Department of Surgery, Saitama Medical Center, Jichi Medical School, Saitama, Japan<sup>5</sup> President, International University of Health and Welfare, Ootawara, Japan**Keywords**

Endoscopic surgical skill; gastroenterology; qualification

**Correspondence**

Taizo Kimura, Department of Surgery,  
Fujinomiya City General Hospital, 3-1 Nishikicyo,  
Fujinomiya 418-0076, Japan.  
Tel: +81 544 27 3151  
Fax: +81 544 23 7232  
Email: fmhosa@lilac.ocn.ne.jp

Received: 26 October 2009; accepted 29  
November 2009

DOI:10.1111/j.1758-5910.2009.00034.x

This article will be presented at the  
AsianAmerican Multispecialty Summit IV  
Laparoscopy & Minimally Invasive Surgery,  
organized by the Society of Laparoendoscopic  
Surgeons, in Honolulu, Hawaii, USA, on 10–13  
February 2010.

**Abstract**

**Introduction:** To reduce the complications of endoscopic surgery, the Japan Society for Endoscopic Surgery formed a committee that established the Endoscopic Surgical Skill Qualification System (ESSQS). Here we report on the methods employed and results obtained with the ESSQS over five years in the field of gastrointestinal surgery.

**Methods:** The first ESSQS review was performed in 2004, and examinations have been conducted once a year since then. Applicants must submit a list of patients on whom they have performed surgery (including complications) and an unedited video showing one of the relevant surgical procedures. To assess the applicants' videos, the judging committee prepared "common criteria" (60 points) and "procedure-specific criteria" (40 points). Assessment of videos was done independently by two judges, and the applicant passed the test if both judges assigned a score of 70 points or more.

**Results:** There have been 1369 applicants, and 641 (46.8%) have been successful. The main problem with this system has been a relatively low rate of agreement between the two judges ( $\kappa$  value: 0.29–0.40). However, the incidence of complications is significantly lower in patients treated by successful applicants ( $4.3 \pm 6.8\%$ ) than in those treated by failed applicants ( $5.6 \pm 8.4\%$ ) ( $P = 0.0096$ ).

**Conclusion:** Although the ESSQS could be further improved, this system promotes training and decreases complications.

**Introduction**

The outcome of surgery is not only influenced by a surgeon's knowledge, but also by his or her technical skills. This is particularly true for endoscopic surgery because it utilizes very specific techniques. Since endoscopic surgery was introduced in Japan, the incidence of complications associated with this type of surgery has increased and this is often related to poor technique.

To reduce the incidence of complications, it is necessary to educate the surgeons performing endoscopic surgery based on both a specific program and technical accreditation. The Japan Society for Endoscopic Surgery (JSES) has organized educational seminars, workshops using animals, and lectures about endoscopic suturing. The JSES

also established the committee that initiated the Endoscopic Surgical Skill Qualification System (ESSQS) in 2001, consisting of members selected from the fields of gastrointestinal surgery, obstetric/gynecologic surgery, urological surgery, respiratory surgery, orthopedic surgery, and pediatric surgery. The committee discussed the basic concepts for an accreditation system and decided on the following four points: (1) an applicant must be a specialist in a certain field and have sufficient experience of endoscopic surgery; (2) technical assessment would be based on viewing an unedited video of surgery performed by the applicant; (3) surgeons should only receive accreditation if their technique is good enough for them to act as instructors; and (4) a judging committee should be established for each surgical field.

The first endoscopic surgical accreditation examination was held for obstetric/gynecologic surgeons in 2003, while surgeons from the fields of gastroenterology, urology, and orthopedics were assessed in 2004 (1). In 2008, accreditation for the field of pediatric surgery commenced. Since then, examinations have been conducted once a year in the above-mentioned fields.

Thus far, a similar third-party accreditation system for endoscopic surgeons has not been established in any other country. Here we report on the methods employed and the results obtained with respect to endoscopic surgical accreditation in the field of gastrointestinal surgery.

## Methods

### Selection of judges

The initial judging committee consisted of 25 experienced laparoscopic surgeons who were recommended by the JSES and mutually reviewed videos submitted by the other surgeons. They were all considered to have sufficient technical skills to act as judges and were divided into six groups to assess operations in six different areas (esophagus, stomach, colon, biliary tract, spleen, and endocrine/others). Later, another 49 judges were added by selection from among surgeons who had passed the accreditation examination with a high score. The names of the judges are kept confidential.

### Requirements for applicants

The five requirements listed in Table 1 must be fulfilled by each applicant. Surgeons who have met these requirements must then submit a list of patients on whom they have performed surgery (including complications) and an unedited video showing one of the relevant surgical procedures (Table 2). If the video submitted does not include suturing and knot-tying, another video that includes these procedures also needs to be submitted.

### Method of assessment

To assess the videos, the judging committee prepared sets of "common criteria" (maximum possible score of 60

**Table 1** Requirements for the applicants

1. Applicants must be specialist surgeons.
2. They must have attended the JSES scientific meeting (3 points), educational seminars (4 points), and/or workshops using animals (3 points) (a total of 12 points is required).
3. They must have performed simple surgical procedures (including cholecystectomy) in at least 50 patients or complex procedures (including colectomy) in at least 20 patients.
4. They must be recommended by two instructors.
5. They must have presented at least three reports on endoscopic surgery at scientific meetings and have published two journal articles.

**Table 2** Procedures that can be submitted

- Cholecystectomy, common bile duct clearance
- Esophagectomy, Nissen's operation, Heller's operation
- Distal gastrectomy
- Sigmoidectomy
- Splenectomy
- Inguinal hernioplasty
- Adrenalectomy, nephrectomy, mastectomy, thyroidectomy

**Table 3** Four categories of common criteria (60 points)

Categories	Points	Criteria
Progress of the operation	16	Smooth conduct of surgery Cooperation with assistants
Development of the operating field	15	Proper positioning of the access ports Clear visualization of the operating field Proper use of retractors and grasping forceps
Operative techniques	19	Proper selection and appropriate use of forceps Proper methods of traction and tissue grasping Appropriate and smooth use of the correct type of energy Dissection along the correct plane Correct identification and proper coagulation or clipping of blood vessels
Suturing and knot-tying	10	Suturing Knot-tying

points, Table 3) and "procedure-specific criteria" (maximum possible score of 40 points, examples in Tables 4 and 5). Two judges assess the videos. The applicant's name was not disclosed to the judges, and vice versa. If both judges assigned a score of 70 points or more, the applicant passed the test. If the two judges disagreed, a final decision was made by either obtaining the opinion of a third judge or by group assessment.

### Inter-rater agreement between judges

Inter-rater agreement between the initial two judges was investigated by calculation of Cohen's weighted kappa values (Table 6).

### Incidence of complications

By reviewing the lists of patients submitted by the applicants, differences in the incidence of surgical complications between the patients of applicants who passed the examination and the patients of applicants who failed it were investigated.

### Statistical analysis

To assess the significance of differences, the unpaired *t*-test or the Mann-Whitney *U*-test was used.

## Results

### Results of assessment by the ESSQS

Over four years, there have been 1369 applicants and 641 (46.8%) have been successful. The success rate has ranged between 40.6% and 50.7%, depending on the year (Table 7).

**Table 4** Procedure-specific criteria for distal gastrectomy (2 points are allotted for each item)

1. Is the port located at an appropriate position?
2. Is the operating field secured by appropriate exclusion of the liver, etc.?
3. Are appropriate grasping forceps that can prevent injury to the stomach, duodenum, and small and large intestines used?
4. Are the tissues grasped with appropriate force?
5. Is the full thickness (all layers) of the gastric wall or intestinal wall grasped with the instruments?
6. Is traction placed on the correct site and in the appropriate direction?
7. Is there any serosal damage?
8. Is there any bleeding caused by rough use of grasping forceps?
9. Is the gastrocolic ligament divided appropriately?
10. Are any measures taken to prevent injury to the large intestine?
11. Is the left gastric vein divided appropriately?
12. Is the left gastric artery divided appropriately?
13. Are any measures taken to prevent injury to the pancreas?
14. Is dissection of the No. 1 lymph nodes done appropriately?
15. Is dissection of the No. 3 lymph nodes done appropriately?
16. Is the extent of lymph node dissection sufficient?
17. Is gastroduodenal anastomosis completed without error?
18. Is there an adequate blood supply to the anastomotic site?
19. Is there excessive tension on the anastomosis?
20. Is the anastomosis shaped correctly?

### Success rate for each surgical category

For each surgical category, the success rate differed slightly. It was high for splenic surgery (55.2%), and low for procedures on the colon (43.6%) (Table 8).

### Inter-rater agreement

In 2004, the overall  $\kappa$  value for agreement in rating the applicants was only 0.31, but it increased to between 0.36

**Table 5** Procedure-specific criteria for cholecystectomy (3 points are allotted for each item)

1. Elevation of the gallbladder (GB) and obtaining good exposure without liver injury
2. Retracting the duodenum and transverse colon to secure a good view of the area around Calot's triangle without adjacent organ injury
3. Exposing the tissues around the cystic duct properly
4. Identifying the cystic artery properly and carefully locating the right hepatic artery
5. Identifying the common bile duct
6. Clipping or ligating the cystic duct properly
7. Transection of the cystic duct at the proper level, so that the remnant duct is an appropriate length
8. Dissecting the liver bed in the correct plane
9. Appropriate control of liver bed bleeding
10. Retrieval of the GB

Additional points awarded according to the degree of difficulty in performing cholecystectomy (10 points)

- Uncomplicated procedure (0 points)
- Soft adhesions that are easy to cut or separate (2–4 points)
- Adhesions that are firm and difficult to cut (6–8 points)
- Acute cholecystitis, thick-walled gallbladder, or bleeding that is hard to control (10 points)

**Table 6** Assessment of inter-rater agreement

Judge A				Subtotal
Cohen's kappa values (κ)				
Judge B	a	b	c	a+b+c
	d	e	f	d+e+f
	g	h	i	g+h+i
Subtotal	a+d+g	b+e+h	c+f+i	n
Po = (a+e+i)/n				
Pe = ((a+b+c) × (a+d+g)+(d+e+f) × (b+e+h)+(g+h+i) × (c+f+i))/n²				
κ = (Po – Pe)/(1 – Pe)				

	Judge A			
	A	B	C	D
Weighted kappa values ( $\kappa_w$ )				
Judge B	A	w = 1	w = 0.67	w = 0.33
	B	w = 0.67	w = 1	w = 0.33
	C	w = 0.33	w = 0.67	w = 1
	D	w = 0	w = 0.33	w = 0.67

$$\kappa_w = (Po(w) - Pe(w))/(1 - Pe(w))$$

**Table 7** Results of assessment by the ESSQS (2004–2008)

Year	No. of applicants	Successful	Success rate (%)
2004	422	214	50.7
2005	269	129	48.0
2006	217	88	40.6
2007	203	90	44.3
2008	258	120	46.5
Total	1369	641	46.8

ESSQS, Endoscopic Surgical Skill Qualification System.

**Table 8** Success rate for each surgical category (2004–2008)

Category	No. of applicants	Successful	Success rate (%)
Biliary tract	576	277	48.1%
Esophagus	75	34	45.3%
Stomach	275	128	46.5%
Colon	358	156	43.6%
Spleen	29	16	55.2%
Endocrine, etc.	56	30	53.6%
Total	1369	641	46.8%

**Table 9** Summary of agreement (Cohen's weighted kappa values) between the initial two judges (2004–2008)

Category	2004	2005	2006	2007	2008
Biliary tract	0.18	0.35	0.29	0.20	0.25
Esophagus	0.36	0.28	0.30	0.62	0.40
Stomach	0.37	0.59	0.34	0.28	0.21
Colon	0.40	0.37	0.40	0.37	0.31
Total	0.31	0.40	0.36	0.38	0.29

**Table 10** Complication rates of successful and failed applicants (2004–2008)

Category	Complication rate (%)		P-value
	Successful applicants	Failed applicants	
Biliary tract	3.2 ± 5.9	3.7 ± 7.7	0.871
Esophagus	11.0 ± 11.8	9.7 ± 11.6	0.6538
Stomach	4.7 ± 5.9	7.6 ± 1.0	0.0284
Colon	4.4 ± 6.4	6.6 ± 7.9	0.0048
Total	4.3 ± 6.8	5.6 ± 8.4	0.0096

and 0.40 between 2005 and 2007. However, inter-rater agreement declined to 0.29 for the 2008 ratings. When the different surgical categories were compared,  $\kappa$  values were low for the biliary tract ( $< 0.29$ ), except in 2005, while the values for colonic surgery were always at an acceptable level (between 0.31 and 0.40) (Table 9).

## Incidence of complications

The overall incidence of complications for all surgeries and the incidence of complications specifically after gastric or colonic surgery were significantly lower in patients treated by successful applicants than in patients treated by failed applicants ( $P = 0.0096$ , 0.0284, and 0.0048, respectively) (Table 10).

## Discussion

The most difficult issue with respect to developing a surgical skills qualification system is how to assess surgical techniques reliably and fairly, especially as surgery has been described as an art rather than a science. However, assessing endoscopic surgery may be easier than assessing open surgery. The entire endoscopic procedure is performed by using a TV monitor and recorded on video, which makes it possible for surgeons and judges to have the same view of the operation and maintain anonymity.

There have been a number of reports about technical assessment methods, including the use of videos or watching live surgery. Winckel *et al.* introduced a structured technical skills assessment form, which employs both procedure-specific checklists and a global rating (global assessment) (2). Martin *et al.* developed a method for objective structured assessment of technical skill, which includes specific checklists and a global rating, as well as a pass/fail judgment (3). For assessment of laparoscopic cholecystectomy, Eubanks *et al.* introduced an objective scoring system that rates surgical techniques by adding points for passing each procedure and subtracting points for each failure (4). In 2005, Vassiliou *et al.* reported their method for global operative assessment of laparoscopic skills, which includes global assessment, a 10-item checklist, and a visual analogue scale for surgical difficulty (5). These methods have been used by surgical trainers when assessing residents or young surgeons, and very high inter-rater agreement between the trainers and repeatability of assessment have been reported.

In order to achieve an objective and reliable assessment of endoscopic surgical techniques, the JSES Judgment Committee for Endoscopic Gastrointestinal Surgery devised both common criteria and procedure-specific criteria before starting the examinations. However, the level of inter-rater agreement between the initial two judges was found to be poor and the overall  $\kappa$  value obtained in 2004 was only 0.31 (Table 9). Although the system allowed review by either a third judge or group decision when the initial two judges could not agree, a higher level of inter-rater agreement was thought to be essential. After the completion of the 2004 examinations, the reasons for the low level of inter-rater agreement were discussed by the Judgment Committee; it was concluded that possible

reasons were that surgical procedures differed between institutions (especially cholecystectomy) and that the common criteria and procedure-specific criteria were imprecise. Accordingly, meetings were held to allow judges to discuss differences in surgical procedures and decide on the permissible variations, especially for cholecystectomy. In addition, the common and procedure-specific criteria have been modified and have become more precise year by year. Tables 3–5 show the criteria used in 2008. After these measures were taken, the overall  $\kappa$  values increased to an acceptable level (between 0.36 and 0.40) from 2005 to 2007 (Table 9). However, inter-rater agreement decreased again to 0.29 in 2008. The Judgment Committee is now investigating the reasons for this decline.

The overall success rate of applicants is fairly low (46.8%) (Tables 7 and 8), but this is not because their technical skills are poor. Rather, it means that the requirements of the Judgment Committee are very strict, and only surgeons whose technique is good enough for them to act as instructors pass the test.

To investigate the reliability of this accreditation system, we reviewed the lists of patients submitted by the applicants and investigated differences in the incidence of surgical complications between patients treated by applicants who passed the examination and patients treated by applicants who failed it. We found that the overall incidence of complications and the incidence of complications after gastric or colonic surgery were significantly lower for patients treated by successful applicants than for patients treated by failed applicants ( $P=0.0096$ ,  $0.0284$ ,

and  $0.0048$ , respectively) (Table 10). However, considering that the definition of complications was imprecise and that complications were reported voluntarily by the applicants (verification was not conducted), these results cannot be regarded as sufficiently reliable. Despite this, the findings appear to suggest the usefulness of our accreditation system.

In conclusion, this endoscopic surgery accreditation system established by the JSES is the first in the world. The system still needs to be modified, but it has already led to an increase in educational seminars and lectures in Japan. Since it appears to contribute to the improvement and standardization of surgical techniques, we think that this accreditation system should be continued.

## References

1. Matsuda T, Ono Y, Terachi T *et al.* The endoscopic surgical skill qualification system in urological laparoscopy: A novel system in Japan. *J Urol* 2006; **176**: 2168–2172.
2. Winkel CP, Reznick RK, Cohen R *et al.* Reliability and construct validity of a structured technical skills assessment form (STSAF). *Am J Surg* 1994; **167**: 423–427.
3. Martin JA, Begehr G, Reznick R *et al.* Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997; **84**: 273–278.
4. Eubanks TR, Clements RH, Pohl D *et al.* An objective scoring system for laparoscopic cholecystectomy. *J Am Coll Surg* 1999; **189**: 566–574.
5. Vassiliou MC, Feldman LS, Andrew CG *et al.* A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg* 2005; **190**: 107–113.