The Endoscopic Surgical Skill Qualification System in Urological Laparoscopy: A Novel System in Japan

Tadashi Matsuda,* Yoshinari Ono, Toshiro Terachi, Seiji Naito, Shiro Baba, Tsuneharu Miki, Yoshihiko Hirao and Akihiko Okuyama

From the Endoscopic Surgical Skill Qualification System Committee, Japanese Urological Association, and Japanese Society of Endourology and ESWL, Japan

Purpose: The Japanese Urological Association, and Japanese Society of Endourology and ESWL have established the Endoscopic Surgical Skill Qualification System in urological laparoscopy. The system consists of a urological section as well as gynecology and general surgery sections. We present details of the Endoscopic Surgical Skill Qualification System and year 1 results.

Materials and Methods: Endoscopic Surgical Skill Qualification System requirements test the ability to complete common laparoscopic surgeries in each field. In urology applicants are required to complete adrenalectomies or nephrectomies appropriately and safely. Applicants should have 2 years of experience with laparoscopic practice, in addition to having completed a 6-year formal urological training program. Also, each surgeon must have performed more than 20 laparoscopic surgeries. According to assessment guidelines applicant skills are assessed by 2 referees who view unedited videotapes showing the entire laparoscopic procedure. To establish these referees 6 expert referees were first selected and 23 were then chosen from 36 referee applicants. Each referee had completed more than 100 laparoscopic surgeries and was chosen after video assessments by the 6 initial expert referees.

Results: Of 5,600 certified urologists in Japan 205 applied to this system in its first year, including 6 expert referees and 36 referee applicants. After video assessments by the referees 136 applicants were certified as having appropriate skills, resulting in a 66% pass rate.

Conclusions: The Endoscopic Surgical Skill Qualification System has just started but it has drawn a lot of attention from the public. We hope that this qualification system will help prevent complications of urological laparoscopic surgeries and promote safer surgical procedures.

Key Words: laparoscopy, urology, process assessment (health care), clinical competence

Since the development of laparoscopic nephrectomy in 1991,[1] laparoscopic surgery has become popular in urology and the number of laparoscopic surgeries has increased dramatically in Japan. However, laparoscopic surgery is technically more difficult than ordinary open surgery and several deaths from laparoscopic surgery have been reported in the newspapers in Japan. Therefore, it is of great interest to the public to ensure the surgical competence of each surgeon.

An evaluation of surgical competence should include an assessment of knowledge, technical skill and judgment.[2] Various methods have been introduced to measure these factors objectively. Knowledge can be measured by examinations, such as those of the American Board of Surgery or Japanese Board of Urology. To obtain objective, reliable and valid assessment of skills several methods have been developed,[3] including motion analysis methods,[4] virtual reality based simulators[5] and objective structured assessments using a living animal or bench model.[6] These examinations may be useful to measure basic or minimal skills, such as in assessments in residents. However, competence in highly complicated procedures such as laparoscopic nephrectomy is difficult to measure by a simple evaluation of basic skills.

Competence has traditionally been based on evaluations of the trainee by senior surgeons. However, this is an indirect measure of skill that suffers from subjectivity and bias.[3] Furthermore, in a field of newly developed skills such as laparoscopy the competence of senior surgeons themselves must be determined. The Society of American Gastrointestinal and Endoscopic Surgeons published guidelines for credentialing qualified surgeons in the performance of general surgical procedures using laparoscopy.[7] They included proctoring applicants for privileges in laparoscopic surgery by a qualified and unbiased staff surgeon experienced with open and laparoscopic surgery together with written reports of applicant competence. However, to our knowledge no medical society has developed a nationwide system to objectively qualify the laparoscopic skills of their members.

The JUA and JSEE have established an especially designed qualification system, the ESSQ System, in urological laparoscopy. The ESSQ System covers not only urology, but also other fields of surgery. The whole system is controlled by the Japan Society for Endoscopic Surgery. The standard requirement of this system for skill qualification is the ability to complete common laparoscopic surgeries in each field

Submitted for publication December 26, 2005.

* Correspondence: Department of Urology and Andrology, Kansai Medical University 2-3-1, Shinmachi, Hirakata, Osaka 573-1191, Japan (telephone: 81-72-804-0101; FAX: 81-6-6993-7757; e-mail: matsudat@takii.kmu.ac.jp).

DOI:10.1016/j.juro.2006.07.034

by the efforts of the applicant. Skill assessment is performed by evaluating an unedited videotape of the surgeon completing the entire procedure in double-blinded fashion. We describe the details of the ESSQ System in urological laparoscopy and report year 1 assessment results.

METHODS

The ESSQ System in urological laparoscopy was established in 2003 according to formal agreements at the general assembly of the JUA and JSEE. The system is controlled by the ESSQ System Committee and the members of this panel were selected by the Board of Directors Committee of the JUA and JSEE. This system was designed to certify urologists who have the capability to complete laparoscopic nephrectomies or adrenalectomies safely and appropriately by their own efforts.

Minimum requirements to apply for skill qualification are experience with 20 or greater laparoscopic adrenalectomies or nephrectomies as chief surgeon, attendance at a laparoscopy training course officially approved by the JUA and JSEE, and more than 2 years of laparoscopic practice after completing a formal 6-year urological training program. Formal application must be sponsored by 2 supervisors who have personal knowledge of applicant laparoscopic surgical skill. Applicants submit unedited videotapes showing whole laparoscopic views of 1 operation (adrenalectomy or nephrectomy) performed alone. On application forms the applicants declare that they performed the operations. The applicant pays about $250 (¥30,000) as a submission fee.

Applicant skill is assessed on the unedited videotapes according to the guidelines established by the Committee (see Appendix). A perfect procedure would score 75 points and 1 to 5 points are deducted if there is a dangerous maneuver. A score of greater than 60 points (80%) is required to pass the assessment. Video assessment is performed by 2 randomly selected referees blinded to the applicant name. If either referee disqualifies the video, the final judgment is made according to a consensus on each video by the Referee Committee. Referees are required to write comments to the applicants describing any inappropriate or dangerous maneuvers shown in the video. About $80 (¥10,000) were paid to each referee for each video assessment.

Referees should have experience with more than 100 cases of laparoscopic adrenal or renal surgery, including adrenalectomies, nephrectomies, pyeloplasties or partial nephrectomies. Six expert referees were first selected by the ESSQ System Committee and established after video assessments by each other. The ESSQ System Committee then recruited applicants for referees from experienced members of the JUA and established 23 referees of 36 candidates after video assessments by the 6 expert referees. Thus, 29 referees were established. Each referee also assessed his own video according to the guidelines.

This system was started in 2004. Applicants were recruited from April to July 2004. Video assessment was performed from the beginning of October to the end of March 2005.

Statistical analyses were performed using 1-way factorial ANOVA using the post hoc, Student t or chi-square test with $p < 0.05$ considered statistically significant. Interrater reliability was analyzed using Cohen's κ value.

RESULTS

In 2004, 205 of the 5,600 urologists certified by the JUA applied to the ESSQ System, including 6 expert referees, 36 referee applicants and 163 general applicants. After video assessments by the 29 referees 136 applicants qualified as having appropriate skills. The pass rate of the skill assessment was 66.3%. The pass rate was 100% for expert referees, 88.9% for referee applicants and 60.1% for general applicants.

The table lists the results of the video assessments of the 163 general applicants by the 29 referees. Videos not considered appropriate for skill assessment included edited videotapes not showing the whole procedure and videos showing partial adrenalectomies. Of the 163 applicants 100 applied with nephrectomy and 63 applied with adrenalectomy videos. A total of 23 nephrectomies were hand assisted and 77 were purely laparoscopic. The pass rates of videos showing adrenalectomy, pure laparoscopic nephrectomy and HALS nephrectomy was 50.8%, 70.1% and 52.2%, respectively. The pass rate for pure laparoscopic nephrectomy was significantly higher than that of adrenalectomy (chi square test $p < 0.05$).

Various inappropriate or dangerous maneuvers were observed by referees in the videos. Of the 69 videos that were finally disqualified the type of dangerous maneuver was inappropriate use of hemostatic instruments in 51%, inappropriate clip application in 43%, insufficient hemostasis in 38%, insufficient surgical fields in 33%, incorrect dissection plane due to a misunderstanding of surgical anatomy in 22% and inappropriate dissection of major arteries in 22%. Furthermore, an assistant had a more major and important role than the applicant during the procedure in 13% of the disqualified videos.

Figure 1 shows the results of the assessment of the 144 videos, which the 2 referees scored. The mean score ± SD on the video assessment was 63.4 ± 7.0 (range 35 to 75) and the mean interexaminer difference was 6.0 ± 5.8 (range 0 to 33). The interrater reliability test to qualify (60 points or greater) or disqualify (less than 60 points) between the 2 referees for the 144 videos resulted in a Cohen κ of 0.22 (95% CI 0.06 to 0.40). Each referee assessed 9.3 of the 144 videos (range 8 to 12). The mean score of the video assessment by each referee was 63.3 ± 2.7 (range 57.9 to 67.8). This difference was not statistically significant on ANOVA. The average of the discrepancy from the score of the other referee on each video was −0.2 ± 3.7 (range 6.9 to −8.1), which was statistically significant (ANOVA $p < 0.0005$, fig. 2). The post hoc test indicated that 8 referees deviated significantly from the other referees with regard to the stringency of the video assessment. Of the 8 referees 4 gave significantly higher scores than the others, whereas 4 gave significantly lower scores than the others. In the 29 referees the average qualify/disqualify agreement rate of each video between the decision of each referee and the final decision was 0.84 ± 0.12 (range 0.6 to 1.0). The average agreement rate in 8 referees with deviation was 0.77 ± 0.11, which was not statistically significantly different from that in the other 21 (0.86 ± 0.12, Student's t test $p = 0.536$).

Video assessment in general applicants in year 1

2-Referee Assessment Results	No. Applicants	No. Certified After Committee Discussion
Qualify/qualify	79	79
Qualify/disqualify	46	17
Disqualify/disqualify	19	1
Qualify/inappropriate	2	1
Inappropriate video	17	0
Totals	163	98

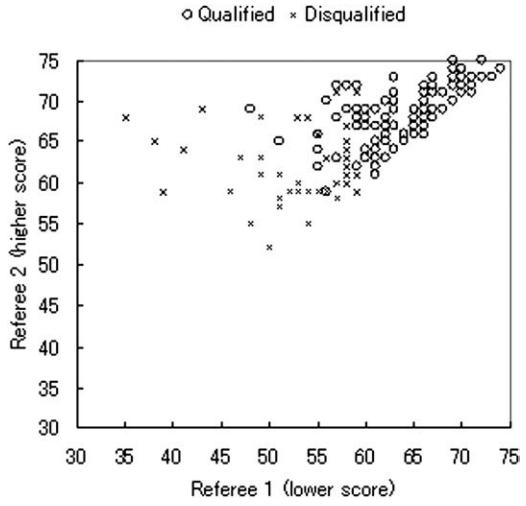

FIG. 1. Results of assessment of 144 videos on applicants scored by each referee.

DISCUSSION

The development of the ESSQ System in urological laparoscopy would serve many purposes. The main goal is to decrease complications due to laparoscopic surgery by evaluating the surgical skills of each applicant and certifying urologists with sufficient skill to perform safe operations. This system also has the potential to set appropriate standards for urological laparoscopic surgery.[8] Feedback on the results of the skill assessment with comments would result in an improvement in the skills of each applicant. Since the assessment is performed in an interinstitutional manner, the system would also promote an equalization of skill levels at different institutions.

In the ESSQ System the assessment of skills is performed on unedited videotapes of the entire laparoscopic procedure with referees blinded to applicants. To evaluate surgical competence intraoperative assessments may be most useful because they evaluate all of the parameters required for a safe operation.[2] The quality assessment of technical skill evaluations should include overall surgical team performance and surgeon leadership of the team as well as individual surgeon performance. The ability to use assistants and communicate in the operating room among other personnel, such as nursing staff and anesthesiologists, and knowledge of specific procedures should also be evaluated. However, assessments in the operation room have several limitations. The number of surgeons established in urological laparoscopy is not sufficient to perform skill assessments in an institution based manner, that is by a physician at the same institution. On-theater skill assessment would be expensive and labor intensive when the examiner comes from elsewhere.[2] Nonblinded, face-to-face assessments may suffer from subjectivity and bias, thus, interfering with an objective and strict evaluation. Assessment from videos without any information on what happened beyond the view of the camera also has limitations to determine real surgeon performance. In our system information regarding patient position, the port sites used by the operator and assistant, the type of scope used, pneumoperitoneum pressure and specific findings during the procedure are written on the application form. Thus, examiner can better understand the circumstances of the surgery.

Referee Committee member selection is one of the most difficult problems requiring resolution to establish a fair system that could be accepted by the members of the society. We could not find a systematic way to select the original expert referees. Six expert referees included of 5 urologists who were selected from different institutional groups and the head of the ESSQ System committee. The expert referees started to perform laparoscopic surgery around 1991 to 1993 as pioneers of urological laparoscopy. After establishing the expert Referee Committee referees were recruited and selected from society members fairly and systematically.

Another major point with respect to assessing surgical skill is to determine what kind of checklist should be used. Several assessment formats have been proposed to evaluate surgical skill. Martin et al developed OSATS using animal and bench models, assessed the surgical skill of residents and reported good reliability and validity for their system.[6] To evaluate live surgery Winckel et al proposed the Structured Technical Skills Assessment Form, which consisted of 2 parts, that is 1) a checklist document of an operation that partitions the procedure into fundamental components and 2) a global rating form of 10 items.[9] In the video assessment system the global rating form proposed in the Structured Technical Skills Assessment Form could not be used because it includes the attitude of the examinee during the operation. We constructed a checklist covering the basic procedures used for laparoscopic adrenalectomies and nephrectomies. Since the main purpose of this system is to decrease complications, we focused on dangerous maneuvers that may result in accidents or complications. The grading system consists of subtracting certain points for each dangerous or insufficient maneuver. Points are deducted if the type of scope, port sites used and patient position resulted in inappropriate sequelae, such as narrow surgical fields. Weight deductions were decided in the guidelines for each inappropriate sequence of events.

Assessment systems should have high reliability and validity. Unfortunately year 1 results of the ESSQ System assessment resulted in rather low interrater reliability. To perform reliable assessment we created assessment guidelines, discussed how to assess the videos in the Referee Committee and performed self-assessments of referee videos. The mean discrepancy between the self-assessment and the assessment by the 6 expert referees was 4.4 with a statistically significant correlation ($r = 0.669$, $p < 0.0001$). The main reason for the low reliability was believed to be the grading system. When dangerous maneuvers are observed several times during the operation, such as improper use of an ultrasonic scalpel, one referee would deduct points only

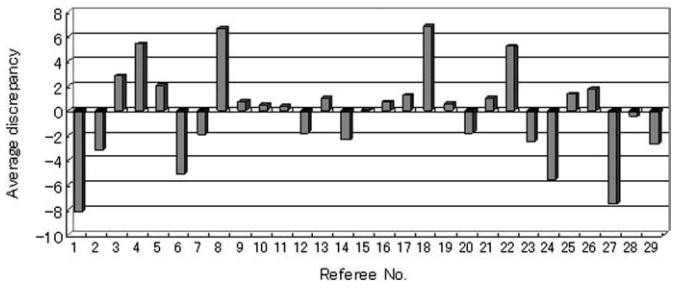

FIG. 2. Average discrepancy from other referee score on each video. Four referees each gave significantly higher and lower scores than other referees.

once, whereas another would deduct a point for each time that the maneuver was performed. Another possible reason for low reliability may be a difference in the stringency of the assessment between referees (fig. 2). Cohen's κ in 72 videos that were assessed only by the other 21 referees, excluding 8 referees with deviation, was 0.42, which was better than the κ of 0.22 obtained in all videos. Further strict guidelines and referee training are required to improve system reliability.

Since so many videos were qualified by 1 referee, while the same videos were disqualified by another, the Referee Committee discussed videos that were disqualified by 1 or 2 referees. The Referee Committee did not discuss the 79 videos that were qualified by each referee. Only 1 of these videos was assessed by 2 of the 4 referees who were defined as nonstringent referees according to the results of the discrepancy evaluation of referee stringency. Although the reliability of referee assessment was low, we believe that the final decision on each video was made appropriately.

Minimum requirements for qualification were a matter of debate. How much experience with adrenalectomy or nephrectomy is necessary to be considered a reliable and safe surgeon? This depends on the learning curve of each surgeon, as determined by talents and learning conditions, but we considered that 20 cases in an average physician after training in basic laparoscopic skills should be the minimum. It is desirable to have an established skill assessment system at a laboratory even for highly complicated procedures such as nephrectomy. The safety of operations performed by surgeons who do not qualify using the ESSQ System is an important matter for patients. According to the guidelines of our society to be an independent chief operator of urological laparoscopic surgery a surgeon must have more than 10 experiences as an assistant of urological laparoscopic surgeries and more than 10 as chief surgeon under the guidance of an established laparoscopic surgeon. After widespread acceptance of the ESSQ System the established surgeon must be a surgeon qualified by this system. This guideline covers not only urology, but also gastrointestinal fields.

We started discussion of this system in 2001 at the JUA and JSEE Executive Committee. It required 2½ years to reach a consensus in our societies and prepare the system. To our knowledge this is the first system to assess the skill of urologists in Japan. Urological laparoscopic surgery is now in crisis in Japan since the reports of severe complications. The consensus in our societies is that a system that ensures the surgical competence of each urologist is mandatory to regain public confidence in urology. This prompted us to develop the ESSQ System in urological laparoscopy. Our final goal is to decrease severe complications due to urological laparoscopic surgery. To prove the predictive validity of this system an outcome survey of operations performed by qualified urologists is mandatory.

ACKNOWLEDGMENTS

Kazuo Suzuki, Hiroaki Mizoguchi and Akihiro Kawauchi assisted with establishing assessment guidelines as members of the 6 expert referees. Eiji Higashihara and Shinichi Oshima, former JSEE Presidents, and Sadao Kamidono, former JUA President, had major roles in establishing this system.

APPENDIX

Video Assessment Guidelines

One point will be deducted if there is an inappropriate maneuver without any sequelae.

Three points will be deducted if there is an inappropriate maneuver with minimum sequelae.

Five points will be deducted if there is an inappropriate maneuver with moderate or severe sequelae.

Assessment Area and Maneuvers

1) Introduction of the ports

Dangerous maneuvers, such as the attachment of the tip of the port to an intra-abdominal organ

Insertion of the port without monitoring

Injury to intra-abdominal organs

2) Maintaining a surgical field

Inappropriate distance between the camera and dissection area

Keeping the surgical field in the center of the camera view

Maintaining an appropriate surgical field with forceps, retractors or hand in hand assisted surgery

Bleeding due to strong traction of the tissues

Performing the procedures under vision with a fogged lens

3) Dissection plane

Entering the mesentery

Entering the perirenal space carelessly

Injury to the psoas muscle

Injury to the adrenal gland

4) Dissection technique

Organ injury during insertion of the forceps

Inappropriate use of the ultrasonic scalpel around the major vessels

Coagulation effects due to electrocautery or the ultrasonic scalpel on tissues where clips have been applied

Coagulation effects due to electrocautery on unintended tissues

Organ injury during dissection procedures

Cutting the tissues without confirming safety

Bleeding due to an inappropriate procedure

Blind use of a hand (HALS)

5) Dissection of the major vessels

Grasping major vessels with forceps roughly

Applying clips halfway on the vessels

Applying clips on other clips

Applying clips not tightly enough

Firing the stapler without confirming the location of its tip

Dangerous dissection of the vessels

6) Control of bleeding

Blind application of clips

Repeat coagulation in an inappropriate manner and location

Not aware of bleeding

No control of bleeding

7) Other

Insufficient bilateral hand coordination

Insufficient coordination of the hand and forceps (HALS)

Performance of important procedures by an assistant

Abbreviations and Acronyms

ESSQ	=	Endoscopic Surgical Skill Qualification
ESWL	=	extracorporeal shock wave lithotripsy
HALS	=	hand assisted laparoscopic surgery
JSEE	=	Japanese Society of Endourology and ESWL
JUA	=	Japanese Urological Association
OSATS	=	Objective Structured Assessment of Technical Skills

REFERENCES

1. Clayman, R. V., Kavoussi, L. R., Soper, N. J., Dierks, S. M., Meretyk, S., Darcy, M. D. et al: Laparoscopic nephrectomy: initial case report. J Urol, **146:** 278, 1991
2. Scott, D. J., Valentine, R. J., Bergen, P. C., Rege, R. V., Laycoc, R., Tesfay, S. T. et al: Evaluating surgical competency with the American Board of Surgery In-Training Examination, skill testing, and intraoperative assessment. Surgery, **128:** 613, 2000
3. Aggarwal, R., Moorthy, K. and Darzi, A.: Laparoscopic skills training and assessment. Br J Surg, **91:** 1549, 2004
4. Smith, C. D., Farrell, T. M., McNatt, S. S. and Metreveli, R. E.: Assessing laparoscopic manipulative skills. Am J Surg, **181:** 547, 2001
5. Gallagher, A. G., Lederman, A. B., McGlade, K., Satava, R. M. and Smith, C. D.: Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance. Surg Endosc, **18:** 660, 2004
6. Martin, J. A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C. et al: Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg, **84:** 273, 1997
7. Society of American Gastrointestinal Endoscopic Surgeons: Granting of privileges for laparoscopic general surgery. Am J Surg, **161**: 324, 1991
8. Reznick, R., Regehr, G., MacRae, H., Martin, J. and McCulloch, W.: Testing technical skill via an innovative "bench station" examination. Am J Surg, **173:** 226, 1997
9. Winckel, C. P., Reznick, R. K., Cohen, R. and Taylor, B.: Reliability and construct validity of a structured technical skills assessment form. Am J Surg, **167:** 423, 1994

EDITORIAL COMMENT

This report represents an extraordinary amount of time and intensive work by these Japanese researchers to develop an especially designed endoscopic surgical skills qualification system in urological laparoscopy. This will undoubtedly become a landmark study and hopefully motivate other collaborative education research groups to perform similar studies for minimally invasive surgery procedures. These will be of benefit to the future of urological surgical training evaluation at the resident and community urologist levels.

In this study 250 urologists were evaluated based on a videotape of a complete laparoscopic renal or adrenal surgical procedure. Based on an unedited videotape evaluation of an entire laparoscopic renal or adrenal procedure participants were assessed by 2 expert referees. Unfortunately the evaluation scoring was based primarily on errors and consideration was not given to the actual steps of the procedure or the performance of these specific steps. The power of this study could have been significantly improved if an OSATS evaluation had been used for the videotape assessment. The OSATS evaluation, which is based on the extensively used and accepted objective structured clinical examination format, is composed of a checklist and a global rating scale. The checklist encompasses specific tasks or steps that must be completed during a technique or procedure. A point is given if the task is performed correctly and no point is given if the task is done incorrectly or not at all. In addition, a qualitative assessment of the procedure is performed using a global rating scale consisting of a Likert-type scale with scores of 1—very poor to 5—excellent. OSATS has been shown to have reliability and validity for assessing operative performance (reference 6 in article). However, the specific cutoff point to distinguish the competent from the incompetent surgeon has yet to be defined for specific procedures. This study would have been an opportunity to create an OSATS score of the competent surgeon for laparoscopic renal surgery. This information then creates the reference for determining a proficiency performance goal that a trainee or resident would be expected to attain.[1] By definition proficiency is advancement in the acquisition of some kind of skill, whereas competence is the capacity to deal adequately with a subject. When setting the proficiency criterion, the experts are used to set the standard but they should reflect a representative sample of the proficiency population and not the top 1% or 5%. If the proficiency criterion level is set too high, trainees will never attain it and, if set too low, an inferior skill set will ultimately be produced. Setting the proficiency guidelines allows resident trainees to have realistic benchmarks by which to gauge their skills acquisition during the training program and by which community urologists can be objectively evaluated in terms of surgical skills. Studies such as this will help create these guidelines and encourage this type of evaluation at the resident and postgraduate training levels in new skills acquisition.

Elspeth M. McDougall
Department of Urology
University of California Irvine Medical Center
Orange, California

1. Gallagher, A. G., Ritter, E. M., Champion, H., Higgins, G., Fried, M. P., Moses, G. et al: Virtual reality simulation for operating room; proficiency-based training as a paradigm shift in surgical skills training. Ann Surg, **241:** 364, 2005