Scientific Skill Assessment of Basic Surgical Dissection and Overall Laparoscopic Performance

Tadashi Matsuda, MD,¹ Kenji Yoshida, MD,¹ Tomonori Habuchi, MD,² Hidefumi Kinoshita, MD,¹ Hiro-omi Kanayama, MD,³ and Toshiro Terachi, MD⁴

Abstract

To improve surgical education, objective and scientific skill assessments are required. There are two types of skill evaluations: assessments of basic surgical skills and assessments of overall surgical performance. To establish a reliable assessment method for surgical dissection, we measured the force applied on the tip of a surgical instrument during dissection of the renal vessels of pigs. The experiments revealed that, during surgical dissection, expert laparoscopic surgeons applied vertical force at the beginning of the stroke and then horizontal force, with minimum vertical force, at the end of the stroke. As an assessment of overall surgical performance, the Endoscopic Surgical Qualification system was developed and has been used for 12 years in Japan. More than 3700 surgeons, including urologists, were determined to have appropriate laparoscopic surgical skills after assessments of unedited videos by referees.

Keywords: skill assessment, laparoscopy, motion analysis

Introduction

SINCE THE DEVELOPMENT of laparoscopic nephrectomy in 1990,¹ laparoscopic surgery has become the surgical treatment of choice for urologic patients primarily due to its decreased invasiveness. Because of the limitations of laparoscopic surgery, including the two-dimensional endoscopic view, no ability to touch the organs, the restricted movement of the laparoscopic instruments without a wrist joint, and unfamiliar anatomy under endoscopic view, laparoscopic procedures suffer from higher technical difficulty compared to open surgery, which results in a steep learning curve. Some patients suffered from complications related to a long operation time and not sufficient techniques of surgeons in the beginning of the laparoscopic era.

To overcome these problems, a variety of educational activities have been conducted worldwide, including dry laboratory, animal laboratory and cadaver training courses, and live surgical demonstrations by hospitals and surgical societies. Development of new surgical instruments such as three-dimensional endoscopes, devices to seal vessels, and surgical robots have also contributed to reducing the difficulty of the procedures. However, new laparoscopic surgeons are still facing numerous challenges, which can feel overwhelming. To achieve a good outcome in laparoscopic

surgery, it is important to assess the surgical skill of each surgeon appropriately and scientifically.²

There are two types of surgical skill assessments: assessments of basic surgical skills and assessments of overall surgical performance, including technical skill, cognition, and judgment. In this study, we describe our findings and experience with these two types of skill assessments.

Basic Surgical Skill Assessment: Force Measurement During Laparoscopic Dissection

Surgical procedures can be divided into several basic components, including dissection, cutting, bleeding control, and suturing. To shorten the learning curve in gaining sufficient skill in laparoscopic surgery, it is required to understand the basics of each surgical component and assess the technique of each doctor.

Surgical dissection is one of the most basic surgical skills. Dissection is carried out by applying force on an instrument. To understand surgical dissection, Yoshida and colleagues measured the force applied on the tip of the dominant-hand instrument during two-handed laparoscopic dissection of the renal vessels from the surrounding fatty tissues, and compared the results between experts and novices.³

¹Department of Urology and Andrology, Kansai Medical University, Osaka, Japan.

²Department of Urology, Akita University Graduate School of Medicine, Akita, Japan.

³Department of Urology, Tokushima University Graduate School of Medicine, Tokushima, Japan.

⁴Department of Urology, Tokai University School of Medicine, Isehara, Japan.

FIG. 1. A force measurement device is incorporated into a spatula-type laparoscopic instrument to measure vertical and horizontal forces applied on the tip of the instrument during surgical dissection.

In the study reported by Yoshida and colleagues,³ a force-measurement device was incorporated into a spatula-type laparoscopic instrument, and vertical and horizontal forces applied on the instrument were measured during each stroke of dissection movement. (Fig. 1) Thirty urologists participated in the study. They were divided into three groups according to their experience with laparoscopic surgery: 10 experts all had more than 80 surgeries (80–750, median = 150) as chief surgeon, 10 experienced surgeons had between 20 and 40 surgeries (median = 27.5), and 10 novices had <5 surgeries (2–5, median = 3.5). Parameters measured in this study include the time to the peak vertical or horizontal force from the start of the stroke divided by the total time (TVF, THF) and the difference between THF and TVF (THF – TVF). (Fig. 2) Ten strokes of the dominant-hand instrument during dissection maneuvers were evaluated for each surgeon.

Yoshida and colleagues demonstrated that TVF for the experts was significantly shorter than for the experienced and novice surgeons, while THF was similar among the groups.³ (Fig. 3) There was a positive correlation between THF-TVF and the number of laparoscopic surgeries for the 30 participants.

FIG. 2. Time to the peak vertical or horizontal force from the start of the stroke during surgical dissection divided by the total time (TVF, THF, respectively) of one stroke and the difference between THF and TVF (THF – TVF) were evaluated.

FIG. 3. Experts showed shorter TVF values than others, while THF was similar among the three groups (* $p < 0.0001$).

This study demonstrated that the novices applied vertical and horizontal force simultaneously during a dissection stroke, while the experts applied vertical force at the beginning of the stroke and then horizontal force, with minimum vertical force, at the end of the stroke. (Fig. 4) Experts changed the direction of force from vertical to horizontal during surgical dissection. This should be the standard technique for surgical dissection. Although this study was performed using a spatula-type instrument, the technique should be similar when a Maryland-type instrument is used. In actual surgery performed by novices, the dissection point slips with little effect on the tissue even though the instrument tip moves on the surface of the tissue. This is caused by a lack of sufficient vertical force at the beginning of the dissection stroke. Furthermore, applying vertical force continuously until the end of the stroke would mean that the dissection would be too deep by the end of the stroke.

Other components of surgery such as knot tying or suturing are also good targets for establishing the most effective and rational technique, and the results could be used to establish more accurate skill assessments for young doctors.

Skill Assessment of Overall Surgical Performance: The Endoscopic Surgical Skill Qualification System

Skill assessment by a preceptor has been done for many years, but it suffers from subjectivity and bias. Objective

FIG. 4. Novices applied vertical and horizontal forces simultaneously during a dissection stroke, while experts applied vertical force at the beginning of the stroke and then applied horizontal force, with minimum vertical force, at the end of the stroke.

FIG. 5. The overall organization of the ESSQ System in Japan, covering many surgical subspecialties. ESSQ=Endoscopic Surgical Skill Qualification.

assessment methods using detailed formats have been proposed, such as the Objective Structured Assessment of Technical Skills (OSATS).⁴ When laparoscopy spread rapidly and widely, a skill assessment system was required to evaluate all surgeons who started to perform laparoscopic surgeries, including senior surgeons since laparoscopy was a completely new approach with new skills required.

In Japan, a new system, called the Endoscopic Surgical Skill Qualification (ESSQ) System, was developed in 2004 by the Japanese Society of Endoscopic Surgery together with subspecialty societies such as those focused on urology, orthopedics, and gynecology.⁵ It is an interinstitutional, nationwide assessment system covering many surgical subspecialties. While it has been built into the educational system of many surgical societies, it is not part of the government certification system.

Skill assessment must be objective, reliable, fair, and valid. To meet these requirements, the ESSQ System set criteria for qualification as follows: (1) applicants are required to have demonstrated an ability to complete common laparoscopic surgeries in different fields with high quality and safety. (2) Skill assessment is performed according to the assessment guidelines in a double-blinded manner by two randomly selected referees using an unedited video recording showing the whole procedure. When the opinions of the two referees do not match, a third referee or the committee makes the final decision.

In 2016, the overall organization of the system, including each subspecialty, is shown in Figure 5. A total of more than 300 referees are working in the system. From 2004 to 2015, a total of 8139 surgeons applied and 3743 were qualified after video assessment, resulting in a pass rate of 46.0% (Fig. 6). In urology, 2303 applied and 1331 were qualified (pass rate = 57.8%).

This system is fair because assessments and decisions are made in a double-blinded manner; only the society clerk knows the identities of the candidate and the referees. It is as objective as possible because the assessments are made according to the assessment guidelines. So, the next question is the reliability of the assessments. To study the reliability of the system, 2440 video assessment scores from 1220 videos assessed by 42 referees from 2004 to 2011 were evaluated by Matsuda and colleagues.⁶ The average discrepancy of the video assessment scores by the two referees for each video was 6.2 points (full score was set at 75 points and ≥ 60 points was required to pass). The pass/fail agreement rate for the two

referees was 68.6% for the 1220 videos. The average number of videos assessed by each referee was 73 for 29 referees working from the beginning of the assessment program, and 25 for the 13 referees who joined the referee committee in 2009. The average assessment score was significantly higher for two referees (both working from 2004) and lower for three referees (two working from 2004 and one from 2009) compared to the overall group. The agreement rate of the results of video assessment by each referee with the final decision of the committee showed a positive correlation, with the number of videos assessed by each referee until 2011, indicating that the referees become more accurate with experience. These results demonstrated that pass/fail assessments of videos showing the entire surgical procedure resulted in moderate reliability of skill assessments, even though the assessments were performed using well-defined assessment guidelines.

The final question is the validity of the system. To study the predictive validity of the system, 20 consecutive laparoscopic case logs attended by 130 urologists, who were qualified in 2004, were prospectively recorded in 2009, and a total of 2590 urologic laparoscopic surgeries were evaluated by Habuchi and coworkers.⁷ Of the 2590 operations, 47.9% were conducted by the qualified surgeon as the chief surgeon, while the other operations were performed by novice surgeons and the qualified surgeon attended the operation as a preceptor. The average operation times for the procedures,

FIG. 6. From 2004 to 2015, a total of 8139 surgeons applied for qualification in the ESSQ System in Japan and 3743 were qualified after video assessments, resulting in a pass rate of 46.0%.

which included laparoscopic adrenalectomy, nephrectomy, prostatectomy, and others, were a little longer than the time reported in the literature from high-volume centers. However, open conversion and allogenic transfusion rates for the 2590 operations were 2.5% and 1.6%, respectively, and major intraoperative (Satava $\geq$ grade 2) or postoperative (Clavien-Dindo $>$ grade 3) complication rates were 1.2% and 0.9%, respectively. These figures were similar to or better than those reported in the literature. When the 1235 operations performed by the qualified doctors themselves and the 1271 operations supervised by the qualified doctors were compared, there were no significant differences in the open conversion, allogenic transfusion, and major intraoperative and postoperative complication rates, indicating that the qualified doctors were effectively supervising the novice doctors.

Discussion

Basic surgical skills have been evaluated using simulators, such as dry laboratory and animal laboratory, and augmented reality simulators. However, it is difficult to assess skills objectively, scientifically, and meaningfully. OSATS,⁴ crowd sourcing⁸ and motion analyses⁹ have been proposed as objective methods. Motion analysis is a promising tool to evaluate the skill of surgeons. Chmarra and associates evaluated the movement of the instrument tip from one point to another and found that all surgeons retracted the instrument first (the retracting phase) and then moved the instrument deep to another point (the seeking phase).¹⁰ Novices showed a longer path and took more time in the seeking phase than experts. Egi and associates developed a system called the Hiroshima University Surgical Assessment Device, which consists of optical scale sensors, microencoders, an experimental table, and a monitor.¹¹ These components were connected to a computer and the movement of instrument tips, rotation angles, distance parameters, and time were measured. The system showed good construct validity. Good construct validity was also reported by many authors using virtual reality simulators, including ProMISTM and LapMentorTM.^{9,12} These assessments that measure, for example, the smoothness of instrument movement, time for a task, or the efficacy of pedaling electrocautery, are useful for measuring the basic skills of trainees. However, these assessments seldom provide information to the trainee on how to improve their skills.

Ideally, skill assessments can also be used to teach the trainee, as the preceptor can communicate positive and negative feedback after the assessment. To make assessments of basic surgical skill more meaningful and useful, the ideal technique for each component of the surgery must be clearly understood. By measuring the force on the tip of the instrument, Yoshida and colleagues found that applying a vertical force first, followed by a horizontal force with minimal vertical force, was an effective and safe method for surgical dissection.³ It should be useful to teach novices to apply vertical force at the beginning of dissection. Furthermore, skill training would be improved if this best practice be incorporated into future systems that measure basic surgical skills. If the best practices for other key components of surgical performance such as knot tying or suturing are revealed, skill evaluation and training would be further exhaustive.

Skill assessment of overall surgical performance is at a different stage compared to the assessment of basic surgical skills. In addition to the surgeon's technical skills, knowledge of anatomy, ability to assess the situation, and decision-making should be evaluated.¹³ Many methods have been proposed to objectively measure actual surgical performance. Winckel and associates proposed the Structured Technical Skills Assessment Form (STSAF) for skill assessment during actual surgery.¹⁴ They developed a 10-item global rating scale and a detailed task-specific checklist and demonstrated good reliability and validity. OSATS, a system similar to STSAF, was proposed to evaluate basic surgical skills in dry or animal laboratories.⁴ This type of combination of global and task-specific rating scales has been used in other assessment systems such as the Global Operative Assessment of Laparoscopic Skills (GOALS),¹⁵ and OSATS and GOALS reported similar results for the same surgeons.¹⁶

Another method of skill assessment for actual surgery is an error-based system. Seymour and coworkers,¹⁷ and Bonrath and coworkers,¹⁸ demonstrated the usefulness and reliability of error-based assessments. A good correlation between an error-based assessment system and the OSATS global rating scale was also reported.¹⁸

In the ESSQ system, the urology group employs an error-based point-deduction scoring system,⁵ while the gastrointestinal surgery group uses OSATS-type point-addition scoring system.¹⁹ The latter group newly developed a laparoscopic surgery-specific global rating scale and procedure-specific step-by-step checklists.¹⁹ The difference of scoring system may be one of the reasons for the difference of pass rates between the two groups, although detailed analyses have not been conducted.

Aggarwal and associates developed a system of motion analysis of actual surgical procedures using a video-based motion-tracking device, and evaluated the time taken, path length, and number of movements for each hand during laparoscopic cholecystectomy.²⁰ They demonstrated good construct validity of the motion analysis data and a good correlation of the motion analysis with skill assessments using the OSATS global rating scale for experienced surgeons.

The ESSQ System in Japan has been used for 12 years and is still being used nationwide for many surgical subspecialties. The system has played a big role in surgical education in Japan; qualifying under the ESSQ System is an important goal during the surgical training of young surgeons. The major drawback of the ESSQ System is the huge amount of time and effort taken by the referees to evaluate the videos. Almost 1000 surgeons, including urologists, apply for qualification every year. We need a way to reduce the time required for evaluating the submitted videos. Crowd-sourced assessments⁸ may have promise if a crowd-sourced full-length review of actual surgical procedures could be proven to reliably correlate with the assessments of trained referees.

Conclusion

In conclusion, skill assessments of surgeons conducting minimally invasive surgery are required to improve surgical education and spread the use of the technique, while ensuring it is safely and appropriately conducted. Motion analysis of surgical performance is a powerful tool for basic skill assessment. To make the analysis more meaningful and useful,

understanding each component of the surgical procedure, such as surgical dissection, is important. Assessments of the surgeon's overall performance in actual surgery are more complicated. Checklist-based systems such as OSATS have been proposed and showed good validity. The Japanese Society of Endourology and the Japanese Society of Endoscopic Surgery developed the ESSQ System, in which surgical performance is evaluated by reviewing unedited videos. This system has been in place and running for 12 years, playing a major role in the education of Japanese surgeons on laparoscopic techniques.

Author Disclosure Statement

No competing financial interests exist.

References

1. Clayman RV, Kavoussi LR, Soper NJ, et al. Laparoscopic nephrectomy. *N Engl J Med* 1991;324:1370–1371.
2. van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP, Dankelman J. Objective assessment of technical surgical skills. *Br J Surg* 2010;97:972–987.
3. Yoshida K, Kinoshita H, Kuroda Y, Oshiro O, Matsuda T. Analysis of laparoscopic dissection skill by instrument tip force measurement. *Surg Endosc* 2013;27:2193–2200.
4. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84:273–278.
5. 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.
6. Matsuda T, Kanayama H, Ono Y, et al. Reliability of laparoscopic skills assessment on video: 8-Year results of the Endoscopic Surgical Skill Qualification System in Japan. *J Endourol* 2014;28:1374–1378.
7. Habuchi T, Terachi T, Mimata H, et al. Evaluation of 2,590 urological laparoscopic surgeries undertaken by urological surgeons accredited by an endoscopic surgical skill qualification system in urological laparoscopy in Japan. *Surg Endosc* 2012;26:1656–1663.
8. Kowalewski TM, Comstock B, Sweet R, et al. Crowdsourced assessment of technical skills for validation of basic laparoscopic urologic skills tasks. *J Urol* 2016;195:1859–1865.
9. Mason JD, Ansell J, Warren N, Torkington J. Is motion analysis a valid tool for assessing laparoscopic skill? *Surg Endosc* 2013;27:1468–1477.
10. Chmarra MK, Jansen FW, Grimbergen CA, Dankelman J. Retracting and seeking movements during laparoscopic goal-oriented movements. Is the shortest path length optimal? *Surg Endosc* 2008;22:943–949.
11. 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;38:705–710.
12. Matsuda T, McDougall EM, Ono Y, et al. Positive correlation between motion analysis data on the LapMentor™ virtual reality laparoscopic surgical simulator and the results from video tape assessment of real laparoscopic surgeries. *J Endourol* 2012;26:1506–1511.
13. Scott DJ, Valentine RJ, Bergen PC, et al. Evaluating surgical competency with the American Board of Surgery In-Training Examination, skill testing, and intraoperative assessment. *Surgery* 2000;128:613–622.
14. Winkel CP, Reznick RK, Cohen R, Taylor B. Reliability and construct validity of a structured technical skills assessment form. *Am J Surg* 1994;167:423–427.
15. 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.
16. Steigerwald SN, Park J, Hardy KM, Gillman L, Vergis AS. Establishing the concurrent validity of general and technique-specific skills assessments in surgical education. *Am J Surg* 2016;211:268–273.
17. Seymour NE, Gallagher AG, Roman SA, et al. Analysis of errors in laparoscopic surgical procedures. *Surg Endosc* 2004;18:592–595.
18. Bonrath EM, Zevin B, Dedy NJ, Grantcharov TP. Error rating tool to identify and analyse technical errors and events in laparoscopic surgery. *Br J Surg* 2013;100:1080–1088.
19. Kimura T, Mori T, Konishi F, Kitajima M. Endoscopic surgical skill qualification system in Japan: Five years of experience in the gastrointestinal field. *Asian J Endosc Surg* 2010;3:66–70.
20. Aggarwal R, Grantcharov T, Moorthy K, et al. An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. *Ann Surg* 2007;245:992–999.

Address correspondence to:

Tadashi Matsuda, MD

Department of Urology and Andrology

Kansai Medical University

Shinmachi 2-5-1, Hirakata

Osaka 573-1010

Japan

E-mail: matsudat@takii.kmu.ac.jp

Abbreviations Used

ESSQ = Endoscopic Surgical Skill Qualification
 GOALS = Global Operative Assessment of
 Laparoscopic Skills
 OSATS = Objective Structured Assessment of Technical
 Skills
 STSAF = Structured Technical Skills
 Assessment Form
 THF = the time to the peak horizontal force
 from the start of the stroke divided
 by the total time
 TVF = the time to the peak vertical force from the
 start of the stroke divided by the total time