学 位 論 文

Development of safe and efficient training system for laparoscopic inguinal hernia repair surgery

(安全かつ効率的な腹腔鏡下鼠径ヘルニア修復術のトレーニングシステムの構築)

2017 年 3 月

北海道大学

サシーム パウデル

Saseem Poudel

学 位 論 文

Development of safe and efficient training system for laparoscopic inguinal hernia repair surgery

(安全かつ効率的な腹腔鏡下鼠径ヘルニア修復術のトレーニングシステムの構築)

2017 年 3 月

北海道大学

サシーム パウデル

Saseem Poudel

Table of Contents

List of Publications and presentations:

A part of this research has been published as following journal papers:

1. Saseem Poudel, Yo Kurashima, Yo Kawarada, Yusuke Watanabe, Yoshihiro Murakami, Yoshiyuki Matsumura, Hiroaki Kato, Kyosuke Miyazaki, Toshiaki Shichinohe, Satoshi Hirano
   Development and validation of a checklist for assessing recorded performance of laparoscopic inguinal hernia repair.
   American Journal of Surgery, 2016 Sep;212(3):468-474. Epub 2015 Dec 13.
   doi: 10.1016/j.amjsurg.2015.09.014.

A part of this research has been presented in the following International conferences:

1. Saseem Poudel, Yo Kurashima, Yo Kawarada, Yusuke Watanabe, Yoshihiro Murakami, Yoshiyuki Matsumura, Hiroaki Kato, Kyosuke Miyazaki, Toshiaki Shichinohe, Satoshi Hirano
   Development and validation of the checklist for the evaluation and feedback of Laparoscopic inguinal hernia repair using surgical video.
   Poster Presentation (English) in Annual meeting of American college of Surgeons October 19-23, 2014 San Fransisco.

2. Saseem Poudel, Yo Kurashima, Yo Kawarada, Yoshihiro Murakami, Kimitaka Tanaka, Hiroshi Kawase, Toshiaki Shichinohe, Satoshi Hirano
   Development of the training system for laparoscopic inguinal hernia Repair
   Oral Presentation (English) in The 12th International Conference of Asia Pacific Hernia Society. October 27-28, 2016, Tokyo

A part of this research will be presented in the following conference:

1. Saseem Poudel, Yo Kurashima, Kimitaka Tanaka, Hiroshi Kawase, Yosihiro Murakami, Yo Kawarada, Yoichi Ito, Fumitaka Nakamura, Satoshi Hirano
   Development and validation of Education System for Laparoscopic Inguinal Hernia Surgery (TAPP) procedure: A multicenter prospective randomized controlled trail.
   Annual meeting of Society of American Gastrointestinal Endoscopic Surgeons. March 22-25, 2017, Houston.

Introduction

Surgical Education

For the better part of last century, the basis of surgical education around the world has basically been the apprenticeship model "See one, do one, teach one" introduced by W.S. Halsted in 1904 [1]. Especially in North America, for the past few decades, this field has been undergoing a significant evidence-based transformation through the incorporation of simulation training, objective evaluation of the resident and feedback[2]. One of the main reasons behind this paradigm shift is the introduction of laparoscopic surgery in the field of surgery. The surgeons suddenly had to overcome two-dimensional imaging and fulcrum effect due to need to work with longer instruments and diminished tactile feedback[3]. They needed to develop new sets of skills. The patient safety was obviously an issue. The introduction of laparoscopy is thought to be responsible for increase of common bile duct injury by 5 times during cholecystectomy[4]. In 1999, a report "To Err Is Human" raised a serious question about medical errors by documenting as many as 98,000 deaths per year caused by preventable medical errors[5]. These have led surgical educators and researchers to look into the alternative to Halsted model.

Despite these advances in the North America, the field of surgical education in Japan is still in its infancy and has not caught up with these changes. Surgical trainees are mostly expected to learn from their mentors and with lack of systematic approach to surgical training. Survey done in the hospitals training surgical residents, highlighted the lack of structured program and evaluation system into surgical residency in Japan[6]. The concerns about patient safety is however similar in this side of Pacific too. Especially in the field of laparoscopic surgery, this country has also seen several high-profile cases that was attributed to medical errors[7].

Laparoscopic Inguinal Hernia Surgery

Inguinal hernia repair surgery is one of the most common surgery performed in the world with over 20 million surgeries being performed every year[8]. This is generally one of the first surgeries to be performed by surgical residents. In recent years, with the advent of laparoscopic surgery, laparoscopic inguinal hernia repair (LIHR) is slowly gaining popularity[9,10]. Studies have shown improved clinical outcome with LIHR compared to open hernia repair[11, 12]. However, even in North America or many countries in Europe open hernia

repair is still performed in majority of the cases[9, 10, 13]. While the lack of consensus among the surgeons about the benefit of LIHR seems to be the main reason behind this[13], technical difficulty of the procedure is another factor limiting the growth of LIHR[11, 12]. Japan also has seen exponential rise in this procedure since 2013[14]. However, we have also seen the increase in complication rate. Survey by Japan Society for Endoscopic Surgery in 2013 reported the recurrence rate of 4% with LIHR compared to 1 % recurrence rate with open method[14]. We also have to consider that those replying to this survey were institutes with interest in laparoscopic surgery and considering the fact that patients with recurrence often avoid visiting the same hospital, the number may be much higher. The lack of technical knowledge can be attributed to one of the major factor behind this. Even for experience surgeons, Birkmeyer et al. [15] have demonstrated that the technical skills varied widely and this was reflected as postoperative morbidity. Though the study focused on bariatric surgery, similar situation can be expected for the other fields. These factors lead me to believe that there is a need for a systematic curriculum for LIHR.

Out of the two major techniques of LIHR, there are more cases of totally extra-peritoneal (TEP) repair reported in the literature. However, in Japan, transabdominal preperitoneal (TAPP) procedure is more popular[14]. While previous meta-analysis has not been able to show superiority of one technique over another in regards to the clinical outcomes, TEP is often considered to be more difficult technique of the two with relatively longer operation time and potentially longer learning curve[12, 16, 17]. One of the main advantages that TAPP procedure offers is the better view of anatomical landmarks and inguinal anatomy[18]. These features make TAPP an easier procedure to start for novices. However, higher risk of complications has been reported during the learning curve and in the cases performed in small volume centers without experienced LIHR surgeons[12].

Assessment and Feedback

For novices, assessment of their performance and providing them with the formative feedback plays a vital role in the development of their skills. However, if the assessment differs from one rater to another then it will end up confusing the trainee. Thus it is very important to have a validated assessment tool that does not differ from one rater to another and can provide them with focused formative feedback.

Several procedure-specific assessment tools to assess LIHR performance in clinical setting have been developed in the past such as Global Operative Assessment of Laparoscopic Skills – Groin Hernia (GOALS-GH) and Operative Performance Rating System (OPRS) [19, 20]. The intended uses of these instruments are for direct observations and are meant to be used by raters who are present in operation theater with the trainee [19, 20].

However, only limited number of hospitals is staffed with expert surgeons in LIHR. Direct observation, while feasible in teaching hospitals with a number of dedicated teaching staff, is more difficult to adopt in community hospitals with limited staffs. Besides the bias due to the relationship between the evaluator and the trainee[21], one of the key concerns is the need for trained rater to be present in the operating room [22, 23]. As the procedures that are being evaluated become more complex, there will be fewer experts who can assess these procedures correctly and give appropriate feedback.

Recently, studies have shown the possibility of using recorded performances to evaluate the surgical performance[22, 23]. Japan Society of Endoscopic Surgery, has been using recorded performance to evaluate endoscopic surgeons and credentialing them for past decade[24]. For past few years, there is evaluation system for LIHR too, however this assessment tool has not been validated and is marred by poor interrater reliability. The assessment tools that has been developed so far are yet to be proved to be an effective metrics for the evaluation of recorded TAPP performance.

Moreover, these tools by its nature, cover only the major steps of the procedure and does not go into the detail on each aspect, limiting the feedback it can provide to the trainees. For this there is a need for more detailed evaluation tool based on the steps of the TAPP procedure which could provide the detailed feedback and also act as an anchor for the trainees to understand the procedure better.

Existing educational tools for the training of LIHR

Various simulation based curriculum has been developed and their educational value has been established for the training of LIHR surgery[25, 26]. However, both of these study deals with TEP procedure and uses simulation at the core of its curriculum. Recent survey in Japan found that less than half of the hospitals where residents are trained to be equipped with some sort of simulation facility, most of which is simple dry boxes[6]. Simulation training, while

effective is considered costly[27]. In contrast to North America, residents are often distributed in wide range of affiliated hospitals, so it is not uncommon to see one or two surgical residents in a teaching hospital. While this is ideal environment from the point of view of case exposure to the residents, it is often not practical or cost effective to build simulation training facility in each of these hospitals. Low cost simulator for the training of LIHR do exist. Simulator developed and validated by Kurashima et al. [28] is a low cost and recently have been made commercially available in Japan. Hernia endotrainer is although not validated is another low cost alternative[29]. However, simulator alone without the trained trainer and curriculum is not effective. With more importance given to clinical work and most of the teaching done inside the operation theater during the surgery[6], most of these hospitals will still struggle to find dedicated time and trained staff for simulation training. There is a need for structured educational curriculum that can train the residents without putting the staffs of the hospitals into any additional strain. There is a need for a system where the residents can be trained using the facilities available readily in the hospital, collect surgical video assess and give the trainees appropriate feedback to help them improve their performance.

These reason, led me to focus on the development of a systematic educational tool for the training of LIHR, TAPP procedure for novices and to determine its educational value.

Objective:

The objective of my research was to develop an effective and safe training system for LIHR, TAPP method.

Since the project involved various aspects, I decided to break it into three phases.

Phase1: Development and validation of the Assessment tool

Phase 2: Development of the educational system

Phase 3: Validation of the new educational system

List of Abbreviations

95% CI	95% Confidience Interval
CG	Control Group
GOALS-GH	Global Operative Assessment of Laparoscopic Skills – Groin Hernia
ICC	Intraclass Correlation Coefficients
IG	Intervention Group
IQR	Interquartile Range
IRB	Institutional Review Board
LIHR	Laparoscopic Inguinal Hernia Repair
OPRS	Operative Performance Rating System
RCT	Randomized Controlled Trial
SD	Standard Deviation
TAPP	Transabdominal Preperitoneal
TEP	Totally Extraperitoneal

Chapter 1

Phase 1: Development and validation of the Assessment tool

Introduction:

In the first phase of the project, my objective was to develop an assessment tool to assess the performance of LIHR, TAPP procedure and demonstrate its validity using the video recordings of the surgical performance.

Materials and Methods

Development of the TAPP assessment tool

An experienced TAPP surgeon, with experience of more than 300 TAPP procedures conducted the task analysis of the TAPP procedure existing textbooks, expert videos and his experience to extract the important steps of the TAPP procedure and details of what was needed to be done in each step. The initial list was developed using iterative methodology based on semi-structured interviews and group discussions of the experienced TAPP surgeons, a surgeon educator with the experience in the development of operative performance assessment tool. Three experienced TAPP surgeons, with experience of more than 100 TAPP cases each, one open hernia expert (>4000 cases), one surgeon educator (developer of GOALS-GH) were individually contacted with this preliminary list for their opinion. Open hernia expert was added to the panel for his knowledge of hernia anatomy based on the sheer volume of the cases he had experienced. The surgeon educator was added to the panel based on his experience with the development of assessment tool and knowledge of TAPP procedure.

The list was modified based on their opinion. This process was repeated multiple times until the consensus was reached among all 5 members of the panel. The final list consisted of 24 steps of TAPP forming the TAPP checklist (Fig 1.).

TAPP Checklist

Operator/DVD no: Evaluator :

(Yes =1, No=0)

	Topics	Explanation	
Trochar Insertion			
1	View	Trochar is inserted under visualization	
2	Position	Trochar is properly located	
Incision of peritoneum			
3	Starting point	First incision is made in an appropriate location on the lateral to inguinal canal	
Creation of Dissection Space			
4	Start of Dissection	After the incision, the proper layer is found and dissection is started	
5	Layer of Dissection	The proper dissection layer is maintained	
6	Traction	Traction is changed according to the direction of the dissection	
Parietalization			
7	Safe Dissection	Recognizes and safely dissects Spermatic chord and vessels	
Reduction of Hernia Sac			
8	Reduction	Properly reduces the hernia sac. Hernia sac is completely reduced in direct hernia.	
Extent of Dissection (Following Structures can be visualized)			
9	Median	Pubic Symphysis, Cooper ligament, Rectus Muscle	
10	Ventral	Rectus muscle, Inferior epigastric vessels, transverse muscle	
11	Lateral	Dissection of lateral triangle up to Superior Iliac spine	
12	Dorsal	Spermatic vessels and spermatic chord	
Mesh Deployment			
13	Size	Proper mesh size is selected	
14	Position	Mesh is placed properly	
15	Stretching	Mesh is stretched properly	
16	Fixation	Mesh is fixed properly	
Suturing of the Peritoneum			
17	Needle movement	Proper Suture Bite and Pitch and proper needle handling	
18	Final appearance	No tear in peritoneum, no slacking of suture and no mesh visible	
Overall			
19	Energy Device	No electric damage done to surrounding organs	
20	Bleeding	No bleeding requiring hemostatic procedure	
21	Instrument handling	Bimanual Dexterity, able to grab the peritoneum finely	
22	Tissue handling	Tissues are handled gently	
23	Flow of procedure	Overall flow of procedure is smooth	
24	Operation Time	Operation time is within 90 Mins	
		Total	

Comment :

Fig1: The TAPP checklist

Each item in the checklist was given the score based on the requirement stated in the definition, with score of 1 if it was deemed adequate and 0 if it was deemed inadequate. The total score is the sum of score of all the items in the checklist with maximum score of 24.

Validation of the Assessment Tool

Various academic and community hospitals were asked to provide unedited videos of TAPP procedure performed by surgeons with all levels of experience. The surgeons were not provided with the checklist prior to their operations. Only cases of unilateral indirect hernia less than 3 cm diameter were included in this study. Recurrence cases, strangulated hernia cases were excluded from this study.

The institutes were asked to provide digital copies of the video recording copies from laparoscopic cameras without any personal information that could reveal the identity of the patients. They were burned onto unmarked blank discs and given numeric codes so as to blind the raters to the operator and the institutes from where the videos originated. The raters were allowed to increase the play speed of the video while viewing at their discretion.

The collected surgical videos were categorized into 3 groups based on the experience of the surgeon at the time of performing the procedure. Those with less than 10 TAPP procedures were categorized into the novice group, those with more than 50 cases were categorized into the experienced group, and the rest were categorized into the intermediate group.

Three blinded raters, from different institutes rated these recorded TAPP procedures using the newly developed TAPP checklist and GOALS-GH, a previously validated assessment tool for LIHR. The raters included one surgeon educator who had got an experience with both the TAPP procedures and assessment tools and was involved in the development of this TAPP checklist (Rater 1) and two experienced TAPP surgeons (> 100 TAPP) (Rater 2 and Rater 3). All these raters were given short lectures about the use of GOALS-GH and the checklist.

Statistical Analysis

Inter-rater reliability for 3 raters (Raters 1 – Rater 3) and between all possible combination of two raters for both the TAPP checklist and GOALS-GH was calculated using intraclass correlation coefficients (ICCs). Inter-group comparisons of checklist scores were made using Kruskal-Wallis Test (novice, intermediate, and experienced). Spearman's correlation coefficient was used for the analysis of the correlation between the scores of the checklist and

the GOALS-GH and between the checklist score and case volume. Single measures were used for inter-rater reliability, while the average score of three raters was used for the analyses of inter-group comparisons and the correlation of Checklists scores with GOALS-GH scores and case volume. Data are expressed as Median and interquartile range (IQR). $p< 0.05$ was considered significant in all analyses. All statistical analyses were done using SPSS version 17 (SPSS, Inc., Chicago, IL).

Results

Demographics

Thirty videos from 11 institutes were included in this study. The demographics of the surgeons whose videos were included in this study were given in Table 1.

Table 1. Demographics of the surgeons whose recorded TAPP performances were analyzed (n=30). Data are presented as n (%) and medians (IQR)

	Novice[a] group	**Intermediate[b] Group**	**Experienced[c] Group**
Number of surgeons	15 (50)	11(37)	4(13)
Postgraduate year after medical school	10 (5-15)	8 (4-10)	18 (15.3-20)
Self-reported TAPP experience (case)	3 (1-5)	21 (12 - 26)	90 (67.3-100)

[a] Novice: TAPP<10 cases, [b] Intermediate: TAPP 10-50 cases, [c] Experienced: TAPP>50 cases

Inter-rater Reliability

ICC for 3 rates was high at 0.75 (95%CI 0.60-0.86) for the newly developed checklist. The ICC between two raters varied from 0.69 (95% CI 0.45-0.84) to 0.79 (95%CI 0.6-0.89). With regards to the GOALS-GH, the ICC for 3 raters was moderate at 0.60 (95%CI 0.41-0.77) (Table 2).

Table 2: Inter-rater reliability. Results are presented as Intraclass Correlation Coefficients (95% Confidence Interval).

	TAPP Checklist	GOALS- GH
Rater 1 vs. Rater 2	0.79 (0.60-0.89)	0.68 (0.42-0.83)
Rater 1 vs. Rater 3	0.77 (0.57-0.88)	0.55 (0.24-0.76)
Rater 2 vs. Rater 3	0.69 (0.45-0.84)	0.60 (0.30-0.79)
Over all	0.75 (0.60-0.86)	0.60 (0.41-0.77)

Rater 1: surgon educator, Rater 2, 3: TAPP experienced surgerons (TAPP>100 cases).

Known Group Validity

The median checklist score was 23.2 (IQR 22.7-24.0) for the experienced group, 22.0 (IQR 19.7-23.0) for the intermediate group, and 18.7 (IQR 15.3-20.3) for the novice group. The difference between the scores of three groups was significant (p <0.001) (Fig. 2).

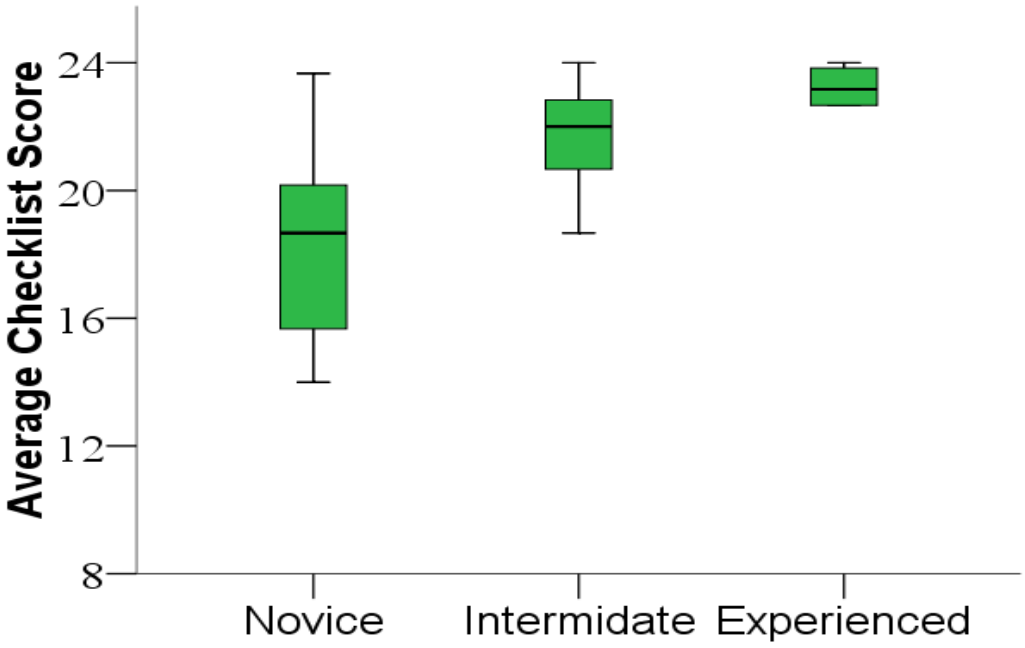


Fig 2: Checklist score and group according to self-reported TAPP experience. Novice: TAPP<10 cases, Intermediate: TAPP 10-50 cases, Experienced: TAPP>50 cases. The horizontal bar represents the median, the box represents the interquartile range (IQR), and the whiskers represent the range.

Concurrent validity

The checklist score showed high correlations with both the case volume and the GOALS-GH score (Fig 3, 4).

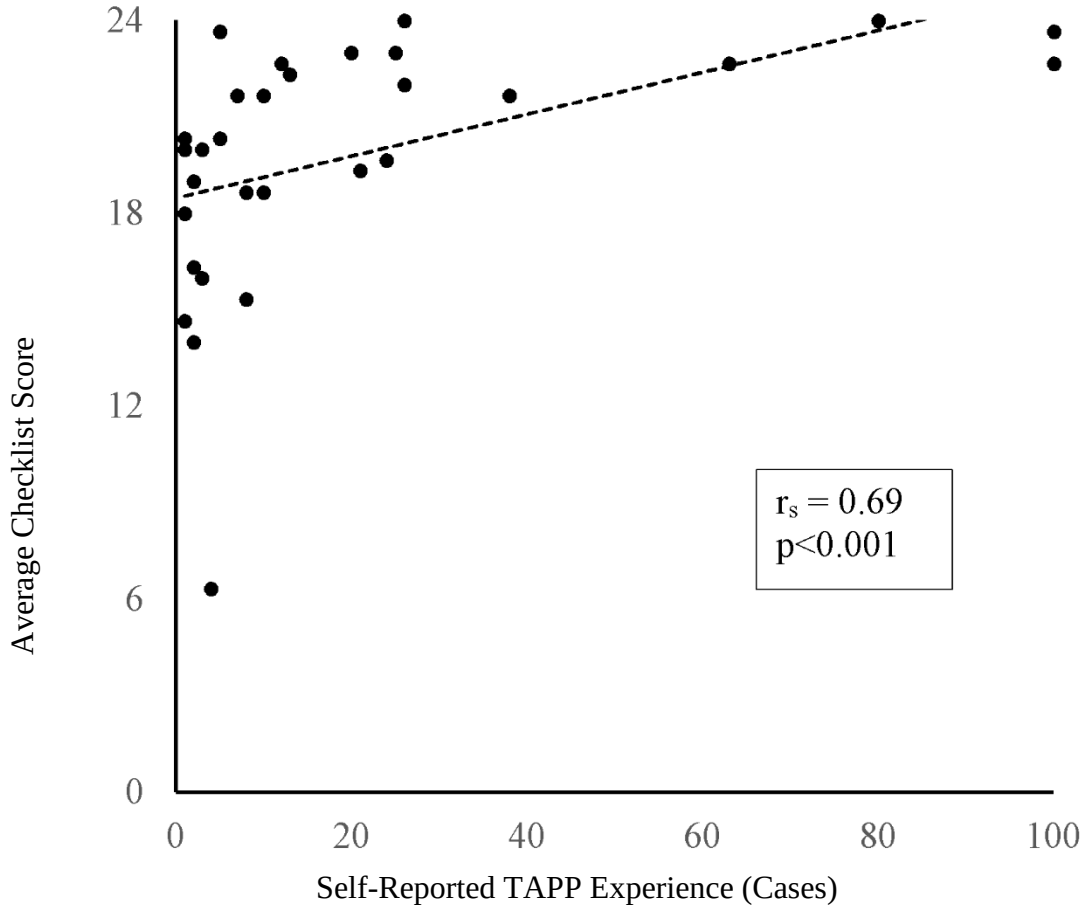


Fig. 3: Correlation between checklist scores and self-reported TAPP experience. Average checklist score represents the average of the checklist score for each case. The correlation is calculated using Spearman Rho (r_s).

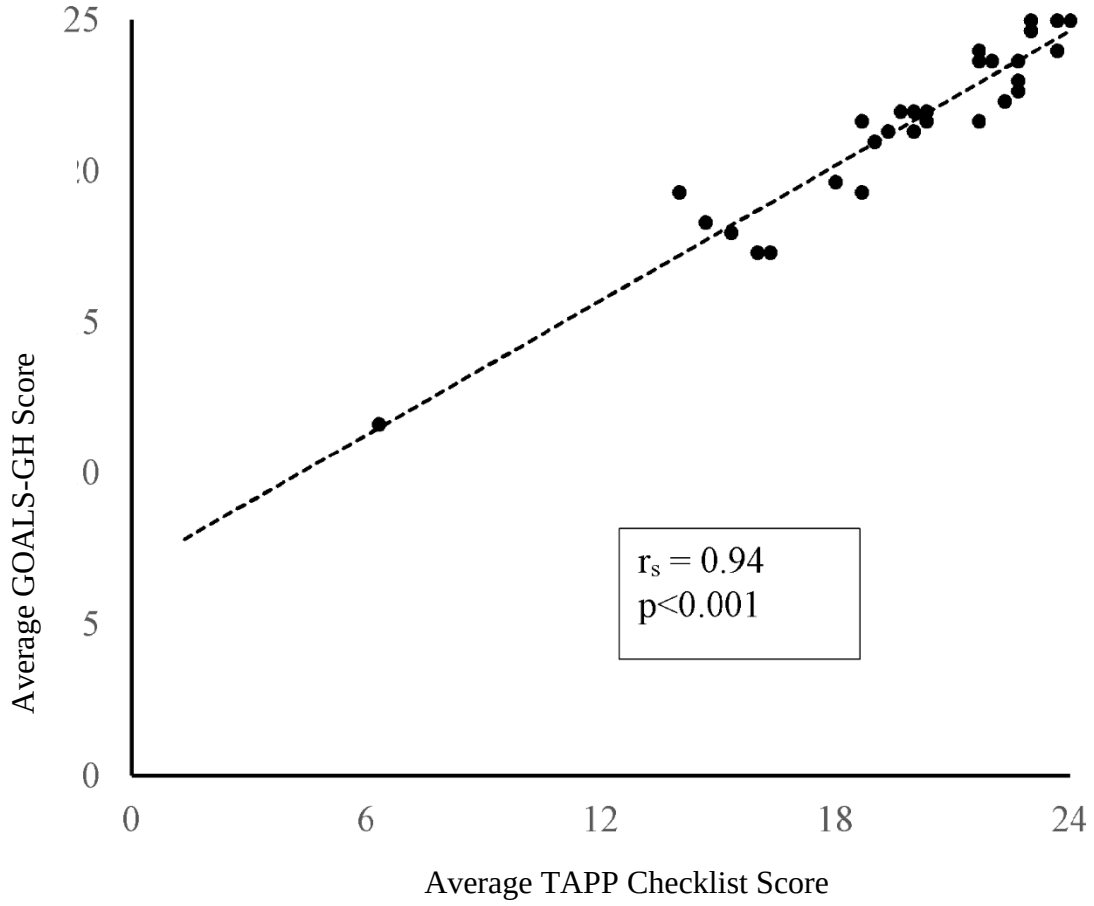


Fig 4: Correlation between checklist scores and GOALS-GH scores. The average Checklist Score and the average GOALS-GH score is the average of the scores given by the 3 raters for each case. Correlation was calculated using Spearman Rho (r_s).

Chapter 2:

Phase 2: Development of the educational system

Introduction

In this phase my objective was to develop the educational tools based on the assessment tool developed in phase 1 and to develop an educational system using the newly developed education tools and assessment and feedback system using the assessment tools developed in phase 1.

Materials and Methods

Development of Educational Tools:

Detailed explanation sheet for the TAPP checklist and the TAPP educational video based on the TAPP checklist was developed as the educational tools.

1. **Detailed explanation for the TAPP checklist**

The TAPP checklist itself is detailed list of the steps of the procedure and simple explanation about what needs to be done at each step. However, due to my effort to keep the checklist within one page, I mainly included the keywords to help the novices in the checklist. A separate detailed explanation sheet was developed explaining each steps in detail. Pitfalls of each step were also explained. For the raters, there was explanation on the criteria to award 1 point or 0 point on each item. For novices, this acted as the information sheet to what needs to be done in each step. TAPP experts and surgical educators were consulted while developing this explanation sheet. International guidelines and local textbooks were also consulted to come up with the final sheet.

2. **TAPP Educational Video**

Two TAPP experts with the experience of more than 200 TAPP cases each were independently contacted and were requested to provide the surgical video clips that would explain each step of the TAPP procedure. They also provided with additional slides explaining the procedure. Residents who had performed less than 10 TAPP procedure were

interviewed on the difficulties they faced when they first performed the TAPP procedures. Surgical video clips and animations were used to explain each item of the TAPP checklist. Anatomical landmarks were explained using surgical videos and animations. The trainees were encouraged to spend time in the dry boxes they had in their hospital to practice their suturing skills and develop their hand eye coordination. Each of these items were arranged separate chapters in the video so the trainees could come back and watch any portion they needed to revise easily.

The educational video was first checked by TAPP experts and surgical educators for any mistakes. It was the piloted to local surgical residents to check if they got the message and found it easy to use. Minor changes were made based on the suggestions of the experts and the residents. Final video was once again checked by the TAPP experts.

Development of the educational system

In addition to the educational tools, a system was created to effectively use these tools for the training of the trainees.

1. **TAPP operators' criteria:**

For this project, I setup the following two criteria that the trainees needed to successfully complete before they were allowed to do TAPP cases on the patients.

a. **Knowledge of the procedure:**
   A simple exam was developed by the researchers to test the trainees of their knowledge of the procedure. The exams contained multiple choice questions related to the TAPP procedure. Total 16 questions were divided into two sets of 8 questions. Care was taken to make both these sets comparable. Both sets of questions were piloted with the surgeons of various levels attending a TAPP educational seminar. Post seminar, the participants had scored full marks. The trainees were required to attain full marks to pass this exam and be eligible to perform the TAPP surgery.

b. **Technical skills:**
   TAPP procedure requires intracorporeal suturing to close the peritoneal flap. The participants were required to be able to complete intracorpreal suture with three knots within two minutes, with the first knot being the surgeons knot. This time was based

on the study by Scott et al. [30], which demonstrated that the residents needed to be trained until they could do intracorporeal sutures with three knots within 111 seconds in order to pass Fundamental of Laparoscopic Skills test.

2. Assessment and feedback

As a part of this education system I incorporated systematic assessments and feedback system for the residents. The residents were assessed and were given feedback by their instructors during the surgery and after the surgery, using validated assessment tools. The residents also did the self-assessment. The portions were the instructor had taken over was given the score of 0. The residents then sent their surgery video to be assessed by the third party along with the short note mentioning the portions were the instructor had taken over the as the surgeon. Blinded TAPP experts who were trained as video raters using the TAPP checklist rated their video and gave their feedback in the comments section. This was sent back to the residents to help them in their forthcoming cases.

For the statistical analysis, higher score was considered as better performance. However, there are multiple ways to obtain the same score. For the clinical and training purposes, all the items were not considered to be equal, and the trainees were asked to concentrate on the items they did not score rather than to be happy with the total score.

Chapter 3:

Phase 3: Validation of the new educational system

Introduction

The objective of this phase was to examine the educational value of the new educational system developed in phase 3 using a multicenter prospective randomized controlled trial.

Materials and Methods

Ethical Issue

The institutional review board (IRB) of Faculty of Medicine at Hokkaido University approved this study (Med 014-0055). Individual IRB approval was also obtained from each of the participating hospitals. The participants were also asked to get written informed consent from the patients informing them that the surgical video of their procedure will be evaluated by experts outside the hospital. The instructors in each institute had full right to determine whether the participants were ready to perform TAPP procedure and take over the procedure if they feel that the participants were having trouble finishing the procedure or if they felt that the patient was at harm. They were allowed to continue with their normal clinical duty and other surgeries during the study period.

Participants

General surgery residents and general surgeons from participating hospitals, who had performed one or less TAPP procedure were recruited for this study from April 2015 to March 2016. The study was designed as a prospective randomized controlled trial (RCT) in accordance with the CONSORT statement (fig 5). The participants were individually explained about this trail and only the participants who provided with the written informed consent were included in this study.

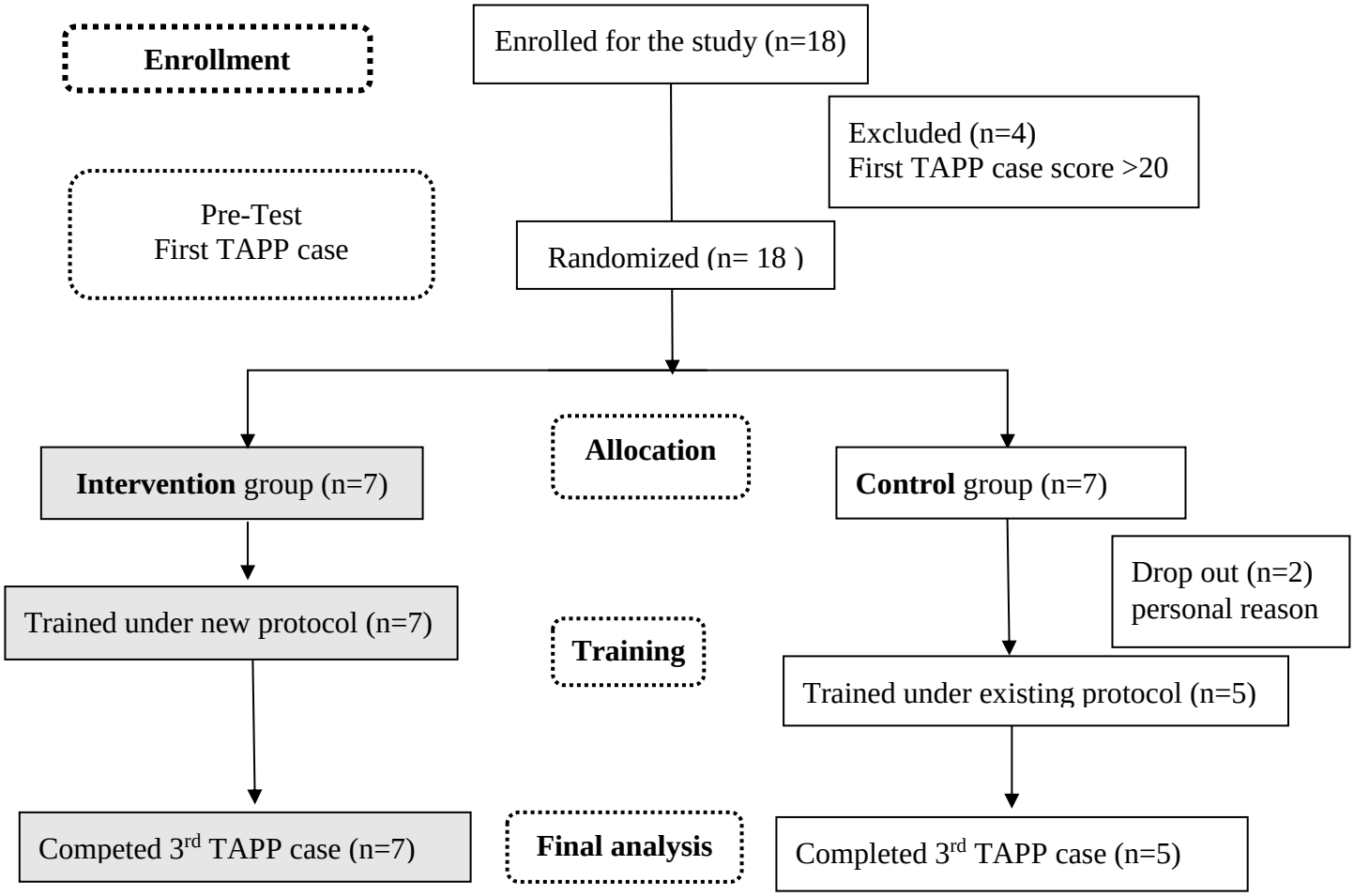


Fig 5: The CONSORT statement

Randomization

Background information on all the participants were collected. The participants were randomly assigned to intervention group (IG) and control group (CG) such that p value between the background data in both groups was >0.4. The participants however were not notified of their grouping until they submitted their first TAPP procedure video.

When there were more than one participants from the same institute, undergoing the study at the same time, participants in the IG were asked not to share their education tools with the participants in other group or the ones who are yet to be notified of their grouping. Instructors in charge of the first case and ones for the CG was also separated from the one who was in charge of the IG during the study period.

Case selection

Institutes were requested to select the male patients with primary indirect hernia, PL2 according to European Hernia Society guidelines as far as possible for the study. The participants were barred from performing surgery on recurrent or incarcerated hernia during the study period.

Assessment of TAPP procedure

All the participants were asked to send the surgical video of their TAPP procedure along with a short note indicating the portions where the instructor had taken over, to the researcher. They were also asked not to perform any more TAPP procedures without the researcher's approval. The videos were checked by the researcher and information that could identify the surgeon or the patients were edited. Digital copies of these video were then randomly rated by one of the three video raters, who had earlier been trained as raters using the TAPP checklist. The raters were blinded to the identity, institute and grouping of the participants.

Exclusion Criteria

The participants were excluded from the study if their first TAPP video checklist score exceeded 20. The participants who had already performed more than one TAPP procedure were also excluded from the study.

Pre-test and Post-test:

All the participants were administered the pre-test and post-test, that was earlier developed and piloted in phase 2. Pre-test was administered after they submitted their first TAPP video. Participants in IG were administered and were required to pass post-test before they were assigned their 2nd TAPP case. The participants in CG were administered post-test after the study period was over and before they were handed out the educational materials.

Training of the Intervention Group

The participants and their instructors were notified on which group they had been assigned to after the pre-test. The participants and the instructors in the IG were provided with the TAPP educational video and the TAPP checklist explanation sheet. They were briefed on how to use the TAPP checklist, the TAPP checklist explanation sheet and the TAPP educational video. The participants were provided with the assessment sheet of their first procedure by the video

raters and were asked to review the educational video and their procedure. They were also asked to practice the intracorporeal suture in the dry box. The researcher tested the participants and were allowed to perform their next TAPP case only after clearing the TAPP operators' criteria that was determined in phase 2. The instructors were instructed to used TAPP checklist to evaluate the trainee after the procedure. They were also encouraged to use the same terms used in the educational video while giving feedback to the trainees. The participants were asked to do the self-evaluation after each procedure using the TAPP checklist and send their TAPP procedure video along with the assessment sheets of both the participant and the instructors. The blinded video rater rated the video using the TAPP checklist and these assessments were sent to the participants to help them with their next procedure.

Training of the Control Group

The participants and the instructors of the CG continued with their regular training and performed TAPP procedure as per the individual hospital criteria. They were asked to submit video of each procedure. After the completion of the study, they were provided with the TAPP checklist, TAPP checklist explanation sheet and the TAPP educational video along with their assessment sheet of their previous surgical procedures.

Outcome

The primary outcome was the total TAPP checklist score of the third TAPP case. I also compared improvement of the score from TAPP case 1 to case 3 in both groups. To look into the different aspect of the TAPP procedure where this educational tool would have impact, I also looked into the difference in the score and the improvement from case 1 to case 3 in each subsection of the TAPP checklist. Subsections, incision of the peritoneum, creation of dissection space, parietalization and reduction of hernia sac were grouped together as dissection of preperitoneal space. The participants were blinded and were not informed of the TAPP case that will be compared as the primary outcome and were asked to provide the videos up to their 4^{th} TAPP case.

The scores given by the blinded video raters were used for all the analysis. If the video provided contained a portion of procedure that was not recoded, the portion was not scored and was handled as the missing data for the total score and the unrecorded sections. The scores of the recorded sections were however used for the analysis of the sections.

Survey:

An anonymous survey was done among the trainees and instructors from affiliated instituted who used this training tool. I used the internet platform provided by Google Forms (Alphabet Inc, California USA) to conduct the survey. The participants and instructors were sent the survey from the platform and were requested to respond to the survey. The survey contained questions assessing the usefulness of the educational tool and the feedback the participant received in 5 point Likert scale with 5 being the best answer. The participants were also queried about the number of times they used the tool during the study period.

Statistical analysis

The primary outcome of this study was the difference between the baseline and performance or third case. Based on my previous data, the mean TAPP checklist score for 1st TAPP case was 17.4±2.7. and for 3rd TAPP case was 18.0. A post training increase in the TAPP checklist score to 20 in IG was considered significant. To demonstrate significant difference between the two groups with and alpha of 0.05 and a beta error of 0.20, the analysis indicated a need for a minimum of five participated in each group.

Student T Test was used to compare the data between two groups. Paired sample test was done to compare the checklist score between TAPP case 1 and TAPP case 3 and test score in pretest and post-test of the same individual. Data are presented as Mean (Standard Deviation).

$p < 0.05$ (two-tailed) was considered significant in all analyses. All statistical analyses were done using SPSS version 17 (IBM, New York, NY).

Results:

Demographics

The eighteen participants from nine institutes were recruited for this study. Four participants (three from the CG and one from IG) were excluded from the study as they met the exclusion criteria. Two participants from CG dropped out of the study due to their personal reasons. Seven participants in IG and five participants in CG completed the study and were included in the final analysis. There was no difference between the backgrounds of the participants. (Table 3)

Table 3: Backgrounds of the participants. Data are presented as Mean (Standard Deviation)

	Control Group (n= 5)	**Intervention Group (n=7)**	**p**
Post graduate year (Year)	3.8 (0.8)	4.5 (2.3)	0.51
Number of TAPP cases as assistant (Cases)	10.8 (10.8)	9.1 (5.8)	0.74
Laparoscopic surgery experience as surgeon (Cases)	17.4 (23.4)	31.3 (33.5)	0.45
Number cases in the institute last year (Cases)	54.0 (11.0)	49.7 (10.1)	0.50
TAPP cases experience of the instructor (Cases)	81.0 (59.2)	84.4 (73.8)	0.93

Primary outcome

In their first case the participants in CG tended to score better 15.4(2.6) compared to the score of 12.4 (3.7) of IG. This difference however was not statistically significant (p=0.16). However, by the third TAPP case, the score of the participants in the IG improved significantly to 21.3 (2.0) (p=0.008) while the change in the score of the participants in CG to 16.4 (4.7) was not statistically significant (p=0.52). The difference between scores between

IG and CG in their third case was also statistically significant (p=0.04). In the analysis of different subdivisions of the procedure, the participants in the IG scored better than their counterparts in CG significantly in important aspects of TAPP procedure; “Dissection of preperitoneal space” (p=0.006) and “Extent of dissection” (0.027). The participants in IG also had better scores than their counterparts in CG in the “Mesh deployment” section, however the difference in the score did not reach the statistical significance with p value of 0.06. In all subsections, the difference between the score in the first TAPP procedure and the third TAPP procedure was not significant for the participants in the CG. However, participants in the IG managed to improve their scores significantly in “Dissection of preperitoneal space”, “Extent of Dissection” and “Mesh Deployment” subsections (Table 4).

Table 4: TAPP checklist scores of the participants. Data are presented as Mean (Standard Deviation).

	Control Group	**Intervention Group**	**p**
TAPP 1 Total Score [24]	15.4 (2.6)	12.4 (3.7)	0.16
TAPP 3 Total Score [24]	16.4 (4.7)	21.3 (2.0)	**0.04***
P	0.52	**0.008***	
TAPP 1 Trocar Score [2]	1.8 (0.4)	1.6 (0.5)	0.45
TAPP 3 Trocar Score [2]	2.0 (0.0)	2.0 (0.0)	1.0
P	0.37	0.175	
TAPP 1 Dissection Score [6]	3.8 (1.1)	3.3 (1.5)	0.53
TAPP 3 Dissection Score [6]	4.4 (0.9)	5.8 (0.4)	**0.006***
P	0.37	**0.021***	
TAPP 1 Extent of dissection Score [4]	2.0 (1.0)	1.9 (1.2)	0.8
TAPP 3 Extent of dissection Score [4]	2.4 (1.7)	4.0 (0.0)	**0.027***
P	0.62	**0.003***	
TAPP 1 Mesh Score [4]	3.4 (0.9)	2.6 (1.3)	0.24
TAPP 3 Mesh Score [4]	3.0 (1.0)	3.9 (0.4)	0.06
P	0. 48	**0.049***	
TAPP 1 Suture Score [2]	1.8 (0.4)	1.7 (0.8)	0.83
TAPP 3 Suture Score [2]	2.0 (0.0)	2.0 (0.0)	1.0
P	0.37	0.36	
TAPP 1 Overall Score [6]	2.6 (1.1)	1.7 (1.1)	0.21
TAPP 3 Overall Score [6]	2.8 (1.9)	3.4 (1.5)	0.54
P	0.83	0.095	

Maximum total possible score in each section is written inside **[ ].** p value in the column represent the difference between the score of the Control Group and the Intervention group. p value in the rows represent the difference between the scores of the first TAPP case and the third TAPP case. * is used to denote the difference that is statistically significant.

Secondary Outcome

In the Pre-test participants from the CG and IG scored 5.8 (1.1) and 5.1 (0.9) respectively. The difference was not statistically significant (p=0.28). In the Post-test, the participants in the CG did not show significant improvement in their score with the score of 6.2 (1.1) (p=0.18). However, the participants in the IG improved their score significantly to 7.8 (0.4) (p<0.001). The Post Test score of IG was significantly better that that of CG (p=0.004). (Table 5)

Table 5: Pre-test and Post-test Scores.

	Control Group	**Intervention Group**	**p**
Pre-test Score	5.8 (1.1)	5.1 (0.9)	0.28
Post-test Score	6.2 (1.1)	7.8 (0.4)	**0.004***
P	0.18	**<0.001***	

Data are presented as Mean (Standard Deviation).
p value in the column represent the difference between the score of the Control Group and the Intervention group. p value in the rows represent the difference between the scores of the Pre-Test and the Post Test. * is used to denote the difference that is statistically significant.

There were no post-operative complications except for one case of seroma in the cases done by the participants.

Survey results

The response rate of the post study survey was 100% for both the instructors (5/5) and the participants (7/7). The participants viewed the TAPP educational video 8.1 (5.2) times and went through the explanation sheet 6.1 (3.7) times on average. All the participants felt that the TAPP educational video was useful in helping them understand the TAPP procedure better. While 6 participants agreed that the explanation sheet was useful while 1 participant was not sure about its educational value.

While all the participants agreed that the TAPP checklist score given by their instructor reflected upon their own level, in regards to the score given by the video raters, 1 participant was not sure about if it reflected his level.

Participants seemed to slightly favor the assessment and feedback by their instructor during the surgery to the feedback from video raters with 6 strongly agreeing and 1 agreeing to the statement that the assessments and feedback from their instructors helped them improve their surgical performance as opposed to 2 strongly agreeing and 5 agreeing to the statement that the assessments and feedback from the video rater helped them improve their surgical performance (fig 6, fig 7).

Fig 6: The result of questionnaire: "Do you agree that assessment and feedback during the surgery was useful in helping you improve your performance?"

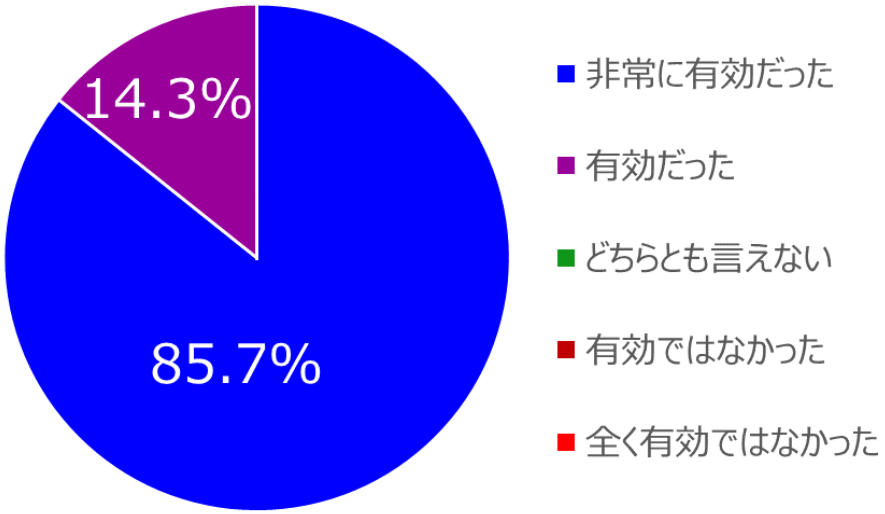
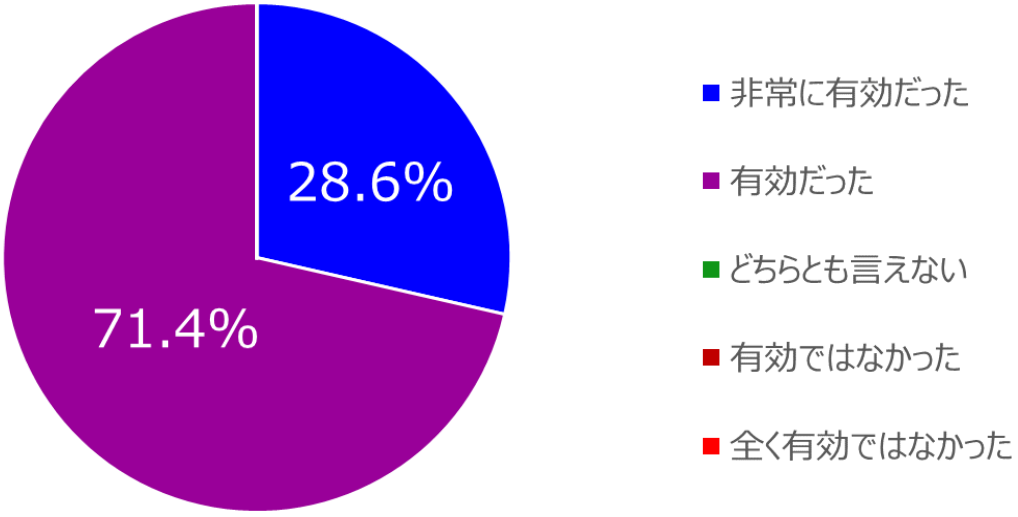


Fig 7: The result of questionnaire: "Do you agree that assessment and feedback by the video rater was useful in helping you improve your performance?"



Sixty percent of the Instructors agreed that both the educational video and the explanation sheet was useful in instruction of the trainees and rest were not sure about its usefulness. Only 40% of the instructors felt that the TAPP checklist was useful in assessing the trainees where was 60% not sure. Eighty percent of the instructors felt that assessment and feedback based on the TAPP checklist was useful in improving the surgical skills of the trainees. However, 60% of the instructors felt that it was taxing for them to use the assessment tools and the opinion was divided with 40% of the instructors wishing to continue, 40% undecided and 20% not wishing to continue using the assessment tools after the study period (fig 8, fig 9).

Fig 8: The result of questionnaire: "Do you think that it was taxing for the instructors to assess the residents and give them feedback in every case?"

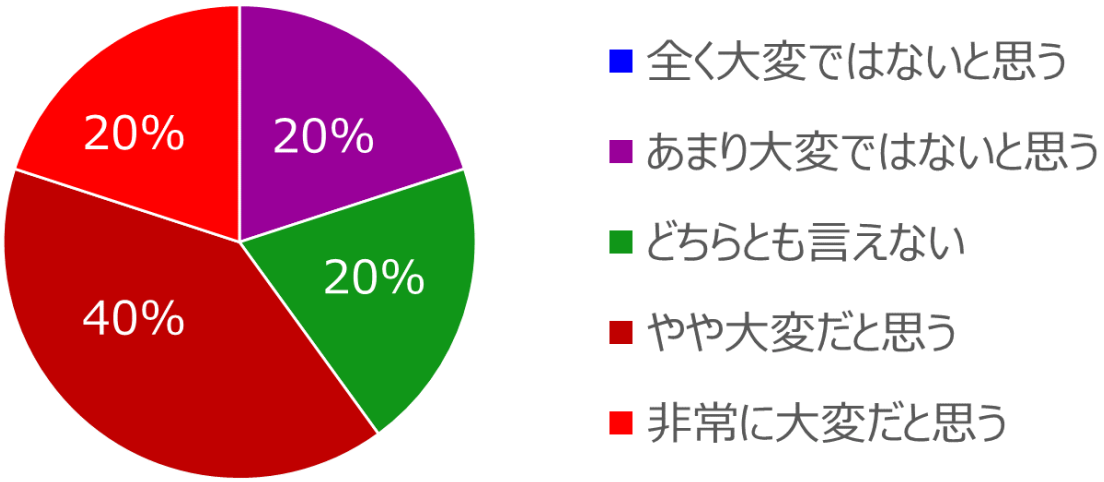
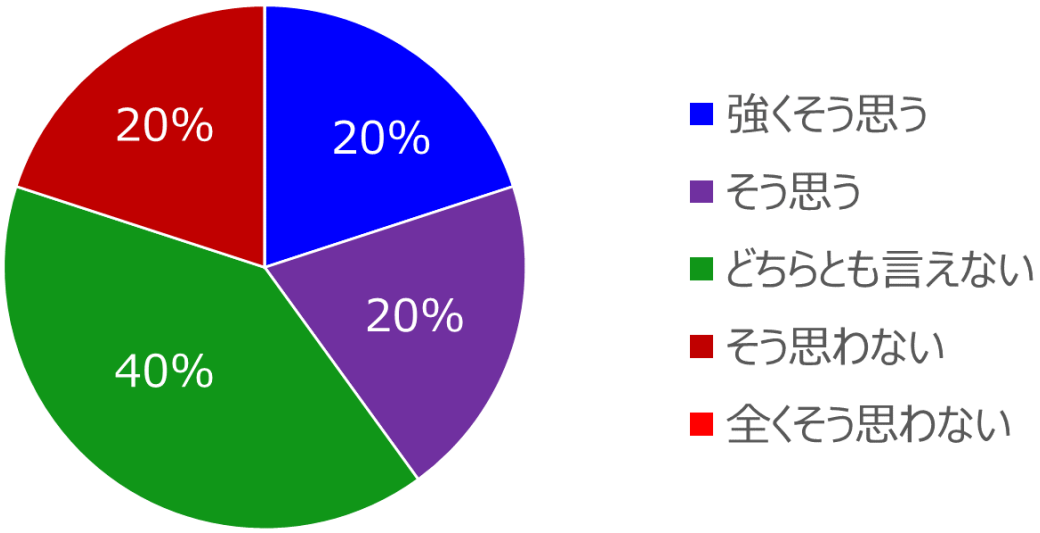


Fig 9: The result of questionnaire: "Will you continue assessing and giving the feedback to your trainees after this research project is over?"



Discussion

In first phase of this project, a novel assessment tool for the TAPP procedure was systematically developed and validated referring to the recommendations given by Ahmed et al.[31]. Task analysis was done to systematically list the important steps and assessment items of the TAPP procedure, and the results were then circulated among experts from multiple centers to attain consensus on the list and its definition, which formed the TAPP checklist. I then validated this tool using 3 blinded raters from different institutes. Although the raters scored recorded performance without a formal rater training prior to assessments, based on the definitions provided in the checklist, ICC was very good at 0.75. After the development and validation of the TAPP checklist I continued this work in phase 2 to systematically develop the education tools based it and developed the educational system using these tools for novices with the remote assessment and feedback at its core. I then demonstrated the educational impact of this new educational tool on the novice surgeons using adequately powered multicenter prospective randomized controlled trial.

Assessment tools

Birkmeyer et al. [15] and more recently Mackenzie et al. [32] have demonstrated that, the surgeons who scored higher when their technical skills were assessed, had the fewer postoperative morbidity. So assessing a surgeon is not only a part of the surgical education for the residents but has a wider role in the society to ensure the safety and quality of the surgery for the general public. Regulatory bodies in few countries already have system in place to assess the surgical procedure and grant credential[24, 33]. More studies are however necessary to develop and validate the assessment tools required for granting licenses for these procedures.

I used video recording of the procedure for the assessment procedure. My choice of pursuing video based assessment was mainly driven by the fact that video based assessment widens the scope of surgical assessments to all types of hospitals and program and even to centers in developing countries as a means of remote education. Any surgeons in any hospital, provided they can record their surgical performance, send their video to the experts and have them assessed. There have already been some attempts to teach and assess basic laparoscopic skills using remote training [34]. Using video ratings this education system already incorporates remote assessment and feedback system. it does not need to set up simulators or other system. This makes it an ideal system which can be adapted to develop a remote training system for LIHR.

Using video recordings for assessments has its share of pitfalls. Raters can evaluate the recorded performance at their convenience, but it also means that they have to give time outside the operating room to evaluate the video recordings. Although Dath et al. [22] have reported decreased time required for evaluations of video recordings compared to direct observation, by allowing the raters to fast forward without affecting the reliability of the assessment, commitment and dedication are needed on the part of the experts to evaluate the videos in their free time[23]. Despite these drawbacks, evaluation using video recordings enables remote evaluations of the procedure and provides an opportunity to novice surgeons to learn about their procedures from experts helping standardize the procedure.

As the complexity of the procedure increases, trainees also need more detailed assessment and feedback. The global rating scales are short and concise, focusing mainly on the general aspects of the procedure. GOALS-GH has only 5 items to assess major steps in the performance of LIHR. While it has been shown to be able to evaluate the trainees both in the simulator setting and in the operating room setting, it would require detailed discussion between the trainer and trainee to get the adequate feedback from the rater[19]. However, the checklist, by its detailed nature, would provide trainees with assessment and feedback on exactly what aspect of their TAPP procedure is inadequate. Gallagher et al. [35] have also concluded that checklists are more suitable as assessment tools as the complexity of the procedure increases.

Using global rating scales, to assess video recordings of the procedures seems to affect the reliability of the tools. Vassiliou et al.[36] and Scott et al.[37] have reported lower ICCs for their assessments when they validated their tools for the use by video raters. The present study showed the ICC between two video raters for GOALS-GH to be between 0.55-0.60, compared to 0.88 between two direct observers reported by Kurashima et al.[19]. With assessment using video recordings, technical aspects are easier to assess than the knowledge of the operator. In their review of the literature, van Hove et al.[38] found that four of five checklists were also used for video rating, whereas just two of the eight global rating scales they reviewed were used for video rating. The present study showed better reliability with checklists compared to a global scale while using video raters. These results demonstrated that the procedural checklist was better option for providing formative assessment using video recordings.

Assessment and Feedback

One of the most important part of this educational system was the incorporation of assessment and feedback system using a valid assessment tools both from their instructors in their hospital and from video rater. These were designed to help the novices understand the aspect they were weak at and focus on that aspect before their next case. Since, the introduction of OSATS[39], assessment using a validated tool and giving the resident an objective feedback has formed the backbone of most of the systematic educational curriculum.

Studies have shown that the feedback plays an essential role in developing surgical skills [40]. Most of the studies agree that immediate feedback to be most effective [40-42]. In this study, the residents highly valued the assessment and feedback they received during the surgery and after the surgery from their instructors and the video rater. They tended value the feedback they received from their instructors immediately after the surgery higher. However, while most of the instructors also felt that assessment and feedback was helpful in the progress of the trainees, 60% of the instructors felt that the process was too taxing for them and only 40% wished to continue the assessment and feedback process after the research period was over. One of the factors behind this may be attributed to the volume of clinical workload on the instructors. As shown by the survey by Kurashima et al.[6] most of the program directors were not satisfied with the amount of time they could dedicate to teach the residents. Even in North America, where much importance has been placed on surgical education, a survey by Jensen et al. [42] revealed that the residents are in fact not receiving as much feedback as they would have liked. A recent study by Soucisse et al.[43] has shown the effectiveness of the feedback after video assessments. This purposed remote assessment and feedback system could ease the burden from these instructors and give the trainees a chance to have their surgical procedure assessed and receive formative feedback to help them improve their skills.

These assessments and formative feedback using a detailed checklist have helped the novices in identifying their weakness. They were encouraged to focus their training and study in those areas. This may have led the novices to deliberate practice of the areas of their weakness. Deliberate practice has been attributed to better performances in the field of games, music, sports [44]. Recently, quite a few studies have looked into its implication on the field of surgical education [45-48]. Hashimoto et al.[46] demonstrated that the quality of the performance of the novices undergoing deliberate practice was higher in simulation training. Other studies also have similar conclusion. Despite the fact that this educational system lacks the tool for technical training other than the dry boxes that is already present in their hospital, novices using this system made significant progress in their TAPP performance. While I did not

record the amount of time they spent on the dry boxes, the participants went through the educational material multiple times, going through the portions they lacked and improvement in various aspects such as dissection, extent of dissection, mesh deployment were evident by the third TAPP surgery.

Educational systems for LIHR

The previous educational system for LIHR were developed and validated in large university hospitals with a number of residents, teaching faculty and equipped with simulation center and proctors [25,26]. The residents first mastered the technique in the simulators before going to operation theater. However, in Japan and most of the countries outside of North America, very few resident teaching hospitals are actually equipped with the simulation center and have trained faculty to train them outside of the operation theater. Most of the teaching still follows the Halsted model where residents are expected to learn from their instructors by watching them in the operation theater. This method however has long learning curve with novices taking up to 60 cases to master LIHR technique [49].

Bokeler et al. [50] after comparing the operative performance of the pioneers of TAPP surgery in their hospital with the novice surgeons recommended systematically introducing the trainees into TAPP procedure by allowing them to do one part at a time and analyzing the videos of the trainees. They also introduced a simple simulator that could be used to train the novices outside the operation room, however both the simulator and the systematic teaching method recommended in the paper were not validated and were described as the expert opinion of their experience of over 15,000 TAPP cases. This educational system incorporates the assessment and feedback of the surgical video of the trainees and I have enhanced it by using the validated tool for this process so that the trainees get the similar assessment and feedback even when the raters change.

It is said that surgery requires 75% of cognitive knowledge and 25% of technical skills [51]. I focused on the cognitive aspect of the surgery in this education tool to help novices understand the TAPP procedure, understand and visualize the terms used by their instructors. The tool was designed to help the novices verbalize the procedure and help them in decision making during the procedure. The survey showed that the trainees highly valued these educational tools. They seemed more biased toward the educational video that visually explained the procedure to them. The instructors, although positive, were not so enthusiastic

as the trainees. These educational tools were targeted for the novice users, and probably did not have the depth that instructors were looking for.

The novice users who trained with this new educational tool made significant improvement with their TAPP performance in short span of 3 cases. They also made significant progress in the important aspect of the procedure as dissection of preperitoneal space, extent of dissection and mesh deployment. Despite of my attempt to match the trainees in both groups, the trainees in the CG had the tendency to score better in their first TAPP case and also in their pre-test. However, the participants in the CG did not show significant improvement in any aspect of their procedure and at the end of the study were outscored by the participants of IG.

Most of the participants from the both the groups scored nearly full marks in items that were relatively easy like insertion of trocar and ones that were purely technical skill like that of suturing of the peritoneal flap. However, in the items where both technical skills and cognitive skills were necessary, like the dissection of the preperitoneal space, extent of dissection, the difference between the CG and IG was evident. In the extent of dissection, all the participant in the IG had scored full marks in their third case. In the earlier study by Kurashima et al. [25] on TEP repair showed similar trend with educational effect most evident on the creation of preperitoneal space and reduction of hernia sac. These steps are actually the most important steps in the TAPP that is directly related to the patient outcome such as relapse, hemorrhage, hernia neuropathy [12].

Simulator training is considered gold standard for the training for surgical skill [2]. In this study, I have been able to demonstrate a considerable educational impact on the novice surgeons without the use of simulation. Ideally, I would have to demonstrate the educational impact of this educational tool by comparing it with established simulator training. However, as mentioned earlier, this educational system is developed for a setting where simulator training is not feasible. I have demonstrated that even in these settings, using systematic approach to training of surgical skills is superior to the traditional training methods. Moreover, I have demonstrated the educational impact of this tool using the trainees distributed in a wide range of affiliated institutes. This adds to the generalizability of my educational system, which is one of the strongest points of this educational system.

Limitations

The present study has some limitations. For the development of the assessment tools, recent studies have been using consensus methodology such as the Delphi method to develop

assessment tools, and this is considered superior due to the anonymity of the participants [33, 52]. I was not able to adopt this method due to technical limitations, but I worked to limit bias among the experts by holding discussions with each expert individually in multiple rounds. Several items in the checklist and its definition consist of "properly" or "appropriate" which makes these items subjective and open to interpretation by the users. This was the result of an attempt to respect the procedural difference that exists between different institutes in non-critical aspects of the procedure like torcher location and my attempt to keep the assessment tool as concise as possible. I have remained objective in the critical aspects of the procedure like extent of the dissection by clearly stating the structures that need to be visible. On the explanation sheet for the checklist for the trainees and raters where each step was further explained in detail so that they could understand what exactly was the meaning of terms "properly" and "appropriate" for each step.

This study has not been able to demonstrate the difference in patient outcome. Zendejas et al.[26], in their study demonstrated less intraoperative and postoperative complications with the trainees who underwent simulation training. This study was not powered to detect the difference between the complications between the two groups. All the participants underwent the procedures under the controlled environment where their instructors took over whenever they felt that the trainees did not know what they were doing or risked the patient safety. While this led to decreased score in the performance of the trainee, it did not necessarily result in the complications in the patients.

Future potential

This educational system can be easily implemented in any hospital to train the trainees with LIHR. It has potential to improve the patient outcome by enhancing the quality of surgery performed by the trainees during the early stages of their learning curve. It also has a potential to serve as template to develop educational tools for other complex surgical procedure.

Summary and Conclusions:

- In the first phase of this project I developed a novel metrics for the assessment of surgical skill related to the TAPP procedure.
- Using recorded TAPP performance I demonstrated that this assessment tool had high interrater reliability, had ability to differentiate between novices, intermediate and experienced TAPP performer and showed high concurrent validity with the number of TAPP cases performed by the surgeon and the GOALS-GH score of the same procedure.
- Using the detailed nature of this assessment tool, I developed the educational tools and educational system with assessment and feedback system at its core. The users of this new educational system found it to be very useful in improving their surgical skill of TAPP procedure.
- In third phase of this project I conducted an adequately powered multicenter prospective randomized controlled trial and demonstrated that the novices from variety of affiliated who underwent the training using this educational system outperformed their counterpart who underwent traditional training method that existed in their hospital.
- This new educational tool has the potential to reduce the learning curve of the novices for TAPP procedure.
- As it does not use any expensive setup, it can be implemented at almost any institute in the world.
- This educational system has potential to improve the patient outcome improving the quality of surgery.
- While I have just concentrated on one surgical procedure in this educational project, there is a need for educational system for other surgical procedures that are commonly performed by the novices to help them shorten their learning curve.
- This educational system that I have developed has a potential to serve as template to develop educational tools for other complex surgical procedure.

Acknowledgements

First and foremost, I would like to thank my mentors, Associate Professor Dr. Yo Kurashima and Professor Satoshi Hirano for encouraging me to undertake this huge project and supporting me in each and every step of my endeavor and making it possible. I am especially grateful to my mentor Dr. Kurashima who not only mentored me in each and every step of my project but also helped my project by lending his expertise in development of the assessment tools and also rated numerous surgical videos during the validation procedure of the assessment tool and the educational system. I am especially indebted to him for introducing me to the world of surgical education research and encouraging me to pursue my dreams.

I would also like to thank Dr. Yo Kawarada, from KKR Tonan Hospital, department of surgery and Dr. Yoshihiro Murakami from Asahikawa city general hospital, department hospital, for lending their expertise the development of the TAPP checklist and later on rating numerous surgical videos to validate the TAPP checklist. I would also like to thank them for providing me with their surgical videos and educational slide for the preparation of the educational video and also supervising the educational video.

I would like to thank Dr. Hiroaki Kato from Teine Keijinkai Hospital, department of surgery, and Dr. Kyosuke Miyazaki from Miyazaki Hernia and Surgery Clinic for their valuable insights and assisting in the development of the TAPP checklist.

I would like to thank Dr. Yoshiyuki Matsumura for agreeing to rate the TAPP videos for the validation process and Dr. Kimitaka Tanaka and Dr. Hiroshi Kawase for rating numerous TAPP videos during the validation of the educational system.

I would like to Dr. Yusuke Watanabe with mentoring me and assisting me in the validation process of the TAPP checklist. I am very grateful to Dr. Yoichi M. Ito from department of biostatistics for mentoring me and providing me valuable insights regarding the statistical analysis of the project.

I would like to thank Dr. Tomoko Mizota, Dr. Mizuna Takahasi for assisting me in the collection of data during my absence. I would also like to thank all the staffs and residents of Hokkaido University Department of Gastroenterological Surgery II, for helping me develop this project by critically analyzing each and every step and encouraging me in every step of this project.

I am grateful to all the instructors and surgeons from the institutes, who agreed to participate in the various stages of this project and made this project a success.

Last but not the least, I would like to thank my parents who showed the importance of education and have always encouraged me to pursue my dreams. Special mention to my dear wife, Bhabana Thapa, a PhD student herself, who understood the importance of research and supported me patiently during these years we have been together.

References:

1 Kerr, B. & O'Leary, J. P. The training of the surgeon: Dr. Halsted's greatest legacy. *Am. Surg.* **65**, 1101-1102 (1999).

2 Tsuda, S., Scott, D., Doyle, J. & Jones, D. B. Surgical skills training and simulation. *Curr. Probl. Surg.* **46**, 271-370 (2009).

3 Gallagher, A. G., McClure, N., McGuigan, J., Ritchie, K. & Sheehy, N. P. An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills. *Endoscopy* **30**, 617-620 (1998).

4 Moore, M. J. & Bennett, C. L. The learning curve for laparoscopic cholecystectomy. The Southern Surgeons Club. *Am. J. Surg.* **170**, 55-59 (1995).

5 in *To Err is Human: Building a Safer Health System* (eds L. T. Kohn, J. M. Corrigan, & M. S. Donaldson) (2000 by the National Academy of Sciences, 2000).

6 Kurashima, Y., Watanabe, Y., Ebihara, Y., Murakami, S., Shichinohe, T. & Hirano, S. Where do we start? The first survey of surgical residency education in Japan. *Am. J. Surg.* **211**, 405-410 (2016).

7 Miyazaki, M. & Otsubo, T. *肝胆膵外科高度技能専門医修練施設腹腔鏡手術緊急実態調査の結果報告について*, <http://www.jshbps.jp/assets/files/files/press_seminar.pdf> (2015).

8 Lofgren, J., Nordin, P., Ibingira, C., Matovu, A., Galiwango, E. & Wladis, A. A Randomized Trial of Low-Cost Mesh in Groin Hernia Repair. *N. Engl. J. Med.* **374**, 146-153, (2016).

9 Zendejas, B., Ramirez, T., Jones, T., Kuchena, A., Martinez, J., Ali, S. M., Lohse, C. M. & Farley, D. R. Trends in the utilization of inguinal hernia repair techniques: a population-based study. *Am. J. Surg.* **203**, 313-317; discussion 317 (2012).

10 Rosenberg, J. & Bay-Nielsen, M. Current status of laparoscopic inguinal hernia repair in Denmark. *Hernia* **12**, 583-587 (2008).

11 Schmedt, C. G., Sauerland, S. & Bittner, R. Comparison of endoscopic procedures vs Lichtenstein and other open mesh techniques for inguinal hernia repair: a meta-analysis of randomized controlled trials. *Surg. Endosc.* **19**, 188-199 (2005).

12 Bittner, R., Arregui, M. E., Bisgaard, T., Dudai, M., Ferzli, G. S., Fitzgibbons, R. J., Fortelny, R. H., Klinge, U., Kockerling, F., Kuhry, E., Kukleta, J., Lomanto, D., Misra, M. C., Montgomery, A., Morales-Conde, S., Reinpold, W., Rosenberg, J., Sauerland, S., Schug-Pass, C., Singh, K., Timoney, M., Weyhe, D. & Chowbey, P. Guidelines for laparoscopic (TAPP) and endoscopic (TEP) treatment of inguinal hernia [International Endohernia Society (IEHS)]. *Surg. Endosc.* **25**, 2773-2843 (2011).

13 Trevisonno, M., Kaneva, P., Watanabe, Y., Fried, G. M., Feldman, L. S., Lebedeva, E. & Vassiliou, M. C. A survey of general surgeons regarding laparoscopic inguinal hernia repair: practice patterns, barriers, and educational needs. *Hernia* (2014).

14 *Japan Society for Endoscopic Surgery*, <http://www.jses.or.jp/>

15 Birkmeyer, J. D., Finks, J. F., O'Reilly, A., Oerline, M., Carlin, A. M., Nunn, A. R., Dimick, J., Banerjee, M. & Birkmeyer, N. J. Surgical skill and complication rates after bariatric surgery. *N. Engl. J. Med.* **369**, 1434-1442 (2013).

16 Bracale, U., Melillo, P., Pignata, G., Di Salvo, E., Rovani, M., Merola, G. & Pecchia, L. Which is the best laparoscopic approach for inguinal hernia repair: TEP or TAPP? A systematic review of the literature with a network meta-analysis. *Surg. Endosc.* **26**, 3355-3366 (2012).

17 McCormack, K., Wake, B. L., Fraser, C., Vale, L., Perez, J. & Grant, A. Transabdominal pre-peritoneal (TAPP) versus totally extraperitoneal (TEP) laparoscopic techniques for inguinal hernia repair: a systematic review. *Hernia* **9**, 109-114 (2005).

18 Castorina, S., Luca, T., Privitera, G. & El-Bernawi, H. An evidence-based approach for laparoscopic inguinal hernia repair: lessons learned from over 1,000 repairs. *Clin. Anat.* **25**, 687-696 (2012).

19 Kurashima, Y., Feldman, L. S., Al-Sabah, S., Kaneva, P. A., Fried, G. M. & Vassiliou, M. C. A tool for training and evaluation of laparoscopic inguinal hernia repair: the Global Operative Assessment Of Laparoscopic Skills-Groin Hernia (GOALS-GH). *Am. J. Surg.* **201**, 54-61 (2011).

20 Larson, J. L., Williams, R. G., Ketchum, J., Boehler, M. L. & Dunnington, G. L. Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents. *Surgery* **138**, 640-647; discussion 647-649 (2005).

21 Vogt, V. Is a resident's score on a videotaped objective structured assessment of technical skills affected by revealing the resident's identity? *Am. J. Obstet. Gynecol.* **189**, 688-691 (2003).

22 Dath, D., Regehr, G., Birch, D., Schlachta, C., Poulin, E., Mamazza, J., Reznick, R. & MacRae, H. M. Toward reliable operative assessment: the reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surg. Endosc.* **18**, 1800-1804 (2004).

23 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.* **247**, 372-379 (2008).

24 Mori, T., Kimura, T. & Kitajima, M. Skill accreditation system for laparoscopic gastroenterologic surgeons in Japan. *Minim. Invasive. Ther. Allied. Technol.* **19**, 18-23 (2010).

25 Kurashima, Y., Feldman, L. S., Kaneva, P. A., Fried, G. M., Bergman, S., Demyttenaere, S. V., Li, C. & Vassiliou, M. C. Simulation-based training improves the operative performance of

totally extraperitoneal (TEP) laparoscopic inguinal hernia repair: a prospective randomized controlled trial. *Surg. Endosc.* **28**, 783-788 (2014).

26 Zendejas, B., Cook, D. A., Bingener, J., Huebner, M., Dunn, W. F., Sarr, M. G. & Farley, D. R. Simulation-based mastery learning improves patient outcomes in laparoscopic inguinal hernia repair: a randomized controlled trial. *Ann. Surg.* **254**, 502-509; discussion 509-511 (2011).

27 Zendejas, B., Wang, A. T., Brydges, R., Hamstra, S. J. & Cook, D. A. Cost: the missing outcome in simulation-based medical education research: a systematic review. *Surgery* **153**, 160-176 (2013).

28 Kurashima, Y., Feldman, L., Al-Sabah, S., Kaneva, P., Fried, G. & Vassiliou, M. A novel low-cost simulator for laparoscopic inguinal hernia repair. *Surg. Innov.* **18**, 171-175 (2011).

29 Jain, M., Tantia, O., Khanna, S., Sen, B. & Sasmal, P. K. Hernia endotrainer: results of training on self-designed hernia trainer box. *J. Laparoendosc. Adv. Surg. Tech. A.* **19**, 535-540 (2009).

30 Scott, D. J., Ritter, E. M., Tesfay, S. T., Pimentel, E. A., Nagji, A. & Fried, G. M. Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training. *Surg. Endosc.* **22**, 1887-1893 (2008).

31 Ahmed, K., Miskovic, D., Darzi, A., Athanasiou, T. & Hanna, G. B. Observational tools for assessment of procedural skills: a systematic review. *Am. J. Surg.* **202**, 469-480.e466 (2011).

32 Mackenzie, H., Ni, M., Miskovic, D., Motson, R. W., Gudgeon, M., Khan, Z., Longman, R., Coleman, M. G. & Hanna, G. B. Clinical validity of consultant technical skills assessment in the English National Training Programme for Laparoscopic Colorectal Surgery. *Br. J. Surg.* **102**, 991-997 (2015).

33 Miskovic, D., Ni, M., Wyles, S. M., Kennedy, R. H., Francis, N. K., Parvaiz, A., Cunningham, C., Rockall, T. A., Gudgeon, A. M., Coleman, M. G., Hanna, G. B. & National Training Programme in Laparoscopic Colorectal Surgery in, E. Is competency assessment at the specialist level achievable? A study for the national training programme in laparoscopic colorectal surgery in England. *Ann. Surg.* **257**, 476-482 (2013).

34 Choy, I., Fecso, A., Kwong, J., Jackson, T. & Okrainec, A. Remote evaluation of laparoscopic performance using the global operative assessment of laparoscopic skills. *Surg. Endosc.* **27**, 378-383 (2012).

35 Gallagher, A. G., O'Sullivan, G. C., Leonard, G., Bunting, B. P. & McGlade, K. J. Objective structured assessment of technical skills and checklist scales reliability compared for high stakes assessments. *ANZ. J. Surg.* **84**, 568-573 (2014).

36 Vassiliou, M. C., Feldman, L. S., Fraser, S. A., Charlebois, P., Chaudhury, P., Stanbridge, D. D. & Fried, G. M. Evaluating intraoperative laparoscopic skill: direct observation versus blinded videotaped performances. *Surg. Innov.* **14**, 211-216 (2007).

37 Scott, D. J., Rege, R. V., Bergen, P. C., Guo, W. A., Laycock, R., Tesfay, S. T., Valentine, R. J. & Jones, D. B. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. *J. Laparoendosc. Adv. Surg. Tech. A.* **10**, 183-190 (2000).

38 van Hove, P. D., Tuijthof, G. J. M., Verdaasdonk, E. G. G., Stassen, L. P. S. & Dankelman, J. Objective assessment of technical surgical skills. *Br. J. Surg.* **97**, 972-987 (2010).

39 Faulkner, H., Regehr, G., Martin, J. & Reznick, R. Validation of an objective structured assessment of technical skill for surgical residents. *Acad. Med.* **71**, 1363-1365 (1996).

40 El Boghdady, M. & Alijani, A. Feedback in surgical education. *Surgeon*, (2016).

41 Garner, M. S., Gusberg, R. J. & Kim, A. W. The positive effect of immediate feedback on medical student education during the surgical clerkship. *J. Surg. Educ.* **71**, 391-397 (2014).

42 Jensen, A. R., Wright, A. S., Kim, S., Horvath, K. D. & Calhoun, K. E. Educational feedback in the operating room: a gap between resident and faculty perceptions. *Am. J. Surg.* **204**, 248-255 (2012).

43 Soucisse, M. L., Boulva, K., Sideris, L., Drolet, P., Morin, M. & Dube, P. Video Coaching as an Efficient Teaching Method for Surgical Residents-A Randomized Controlled Trial. *J. Surg. Educ.* (2016).

44 Macnamara, B. N., Hambrick, D. Z. & Oswald, F. L. Deliberate practice and performance in music, games, sports, education, and professions: a meta-analysis. *Psychol. Sci.* **25**, 1608-1618 (2014).

45 Bhatti, N. I. & Ahmed, A. Improving skills development in residency using a deliberate-practice and learner-centered model. *Laryngoscope* **125 Suppl 8**, S1-14 (2015).

46 Hashimoto, D. A., Sirimanna, P., Gomez, E. D., Beyer-Berjot, L., Ericsson, K. A., Williams, N. N., Darzi, A. & Aggarwal, R. Deliberate practice enhances quality of laparoscopic surgical performance in a randomized controlled trial: from arrested development to expert performance. *Surg. Endosc.* **29**, 3154-3162 (2015).

47 Kirkman, M. A. Deliberate practice, domain-specific expertise, and implications for surgical education in current climes. *J. Surg. Educ.* **70**, 309-317 (2013).

48 Nesbitt, J. C., St Julien, J., Absi, T. S., Ahmad, R. M., Grogan, E. L., Balaguer, J. M., Lambright, E. S., Deppen, S. A., Wu, H. & Putnam, J. B. Tissue-based coronary surgery simulation: medical student deliberate practice can achieve equivalency to senior surgery residents. *J Thorac. Cardiovasc. Surg.* **145**, 1453-1458; discussion 1458-1459 (2013).

49 Choi, Y. Y., Kim, Z. & Hur, K. Y. Learning curve for laparoscopic totally extraperitoneal repair of inguinal hernia. *Can. J. Surg.* **55**, 33-36 (2012).

50 Bokeler, U., Schwarz, J., Bittner, R., Zacheja, S. & Smaxwil, C. Teaching and training in laparoscopic inguinal hernia repair (TAPP): impact of the learning curve on patient outcome. *Surg. Endosc.* **27**, 2886-2893 (2013).

51 Satava, R. M., Gallagher, A. G. & Pellegrini, C. A. Surgical competence and surgical proficiency: definitions, taxonomy, and metrics. *J. Am. Coll. Surg.* **196**, 933-937 (2003).

52 Zevin, B., Bonrath, E. M., Aggarwal, R., Dedy, N. J., Ahmed, N., Grantcharov, T. P. & group, A. Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J. Am. Coll. Surg.* **216**, 955-965 e958; quiz 1029-1031, 1033 (2013).