



Initial surgical performance in robot-assisted radical prostatectomy is associated with clinical outcomes and learning curves

Masahiro Yamazaki^{1,2} · Hiroshi Kawahira³ · Yoshitaka Maeda³ · Kosuke Oiwa⁴ · Hirotaka Yokoyama^{2,5} · Tomohiro Kameda^{2,5} · Jun Kamei² · Toru Sugihara² · Satoshi Ando² · Tetsuya Fujimura²

Received: 10 May 2024 / Accepted: 27 July 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Background The association between surgical performance ratings and clinical outcomes in robotic surgery is poorly understood. Additionally, no studies have reported on the relationship between the surgeon's initial case-skill evaluation and the learning curve in robot-assisted surgery. We evaluated whether an objective surgical technique evaluation score for initial robot-assisted radical prostatectomy (RARP) was associated with clinical outcomes and surgeons' learning curves.

Methods Six surgeons who were trained in and started to perform RARP at our institution were included. Anonymized, unedited videos of each surgeon's 10th RARP case were evaluated by three reviewers, using modified Objective Structured Assessment of Technical Skill (OSATS) scores. We then divided the surgeons into two groups on the basis of these OSATS scores. We retrospectively compared the clinical outcomes and learning curves of the console time of the two groups for consecutive RARPs, performed from March 2018 to July 2023.

Results We analyzed 258 RARPs (43 cases/surgeon), including 129 cases performed by high-OSATS score surgeons (18.2–19.3 points) and 129 cases performed by low-OSATS score surgeons (11.9–16.0 points). Overall, the high-OSATS score group had significantly shorter operation and console times than the low-OSATS score group did (both $P < 0.01$) and their patients' rate of continence recovery by 3 months post-RARP was significantly higher ($P = 0.03$). However, complications, blood loss, and positive margins did not differ between the groups ($P = 0.08$, $P = 0.51$, and $P = 0.90$, respectively). The high-OSATS score group had a significantly shorter console time than the low-OSATS score group did after the 11–20 cases.

Conclusions The OSATS score in early RARP cases can predict subsequent surgical outcomes and surgeons' learning curves.

Keywords Robotic surgery · Surgical skill · Video assessment

Various studies have assessed the quality of surgical performance in different surgical departments [1, 2]. Among these, the Objective Structured Assessment of Technical

Skills (OSATS) score is widely used. Although the OSATS score was originally designed to evaluate residency skills, it is now also used to evaluate laparoscopic surgery [3]. A correlation between OSATS scores and clinical outcomes has been reported in laparoscopic surgery [4], suggesting the importance of assessing surgical performance. However, the quality of robotic surgery has rarely been evaluated using the OSATS score, despite the shift from laparoscopic surgery to robotic surgery [5]. Additionally, most reports on the OSATS scores have assessed the skills of experienced surgeons, with no reports on how the initial skills of novice surgeons affect their subsequent learning curve [4].

Robot-assisted surgery is widely used in urology. In particular, prostate cancer is a major health issue among men worldwide, and opportunities for robot-assisted radical prostatectomy (RARP) are increasing [6]. However, it is unclear if the quality of robotic surgery impacts clinical outcomes.

✉ Hiroshi Kawahira
kawahira@jichi.ac.jp

¹ Department of Urology, Sano Kosei General Hospital, Tochigi, Japan

² Department of Urology, Jichi Medical University, Tochigi, Japan

³ Medical Simulation Center, Jichi Medical University, 3311-1 Yakushiji, Shimotsuke City, Tochigi 329-0498, Japan

⁴ Department of Information and Management Systems Engineering, Nagaoka University of Technology, Niigata, Japan

⁵ Department of Urology, Haga Red Cross Hospital, Tochigi, Japan

Therefore, this study aimed to evaluate the skills of surgeons, who had just started performing RARP, using the OSATS score and to examine whether an objective surgical technique evaluation score for initial RARP was associated with clinical outcomes and surgeons' learning curves. We hypothesized that the OSATS assessment of a surgeon's skill at the beginning of the robotic surgery experience might appropriately predict their subsequent operative performance.

Materials and methods

Study details

This retrospective study included patients who underwent RARP at Jichi Medical University and Haga Red Cross Hospital, Japan, between March 2018 and July 2023. Patients undergoing surgery via a retroperitoneal approach were excluded.

The surgeons whose skills were evaluated were selected on the basis of the following criteria. Surgeons who had started performing RARP at our hospital and had performed over 40 RARP procedures were selected. Japan has a proctoring system that certifies surgeons who have performed more than 40 RARP procedures as instructors. Six surgeons were educated through a mentoring program [7].

Each surgeon's 10th RARP case was assessed using the modified OSATS score, and an anonymous, unedited video of the surgery was reviewed by three urologists, as a previous report suggested that it takes approximately 10 cases to adapt open surgical skills to robotic surgery [8]. If the 10th case was not a typical case and was unsuitable for evaluation, such as a case involving a large prostate or strong intestinal adhesions, the 9th or 11th case was evaluated instead.

Ethics

The study was approved by the institutional review boards of the respective institutions (Jichi Medical University: 23-114 and Haga Red Cross Hospital: 2023_25). Because this was a retrospective study, consent was obtained from patients and the surgeons, whose surgical skills were evaluated, through the opt-out method. Additionally, informed consent was obtained from the reviewers who assigned the OSATS scores.

Surgery

From March 2018 to March 2022, RARPs were performed with the patient in the lithotripsy position, with the head down at a 25-degree angle using the da Vinci Si system (Intuitive Surgical, Sunnyvale, CA, USA). From April 2022

to July 2023, RARPs were conducted with the patient in the supine position with the head down at a 25-degree angle using the da Vinci Xi system (Intuitive Surgical). Our RARP technique was essentially a modified Vattikuti Institute prostatectomy technique [7, 9, 10].

The surgical procedures were divided into eight steps: (1) Dissection of the peritoneum; (2) Dissection of the superficial vein, removal of periprostatic fat, and exposure of the endopelvic fascia; (3) Anterior dissection of the bladder neck; (4) Dissection of the posterior bladder neck, vas deferens, seminal vesicle, and Denonvilliers' fascia; (5) Resection of the lateral pedicle with cavernous nerve sparing; (6) Resection and suturing of the deep dorsal vein; (7) Rocco stitch; and (8) Urethro-vesico anastomosis. For manipulation of the lateral neurovascular bundle, the nerves were preserved in lobes where prostate biopsy showed no cancer in the peripheral zone.

Surgical skill evaluation by the OSATS

Three urologists with experience of more than 60 RARP cases were selected as reviewers to rate unedited, anonymized RARP videos based on video-modified OSATS.

The video-modified OSATS includes five items: "gentleness," "tissue exposure," "instrument handling," "time and motion," and "flow of operation"; each item is scored on a scale of 1–5 points (1 point: novice level, 3 point: intermediate level, and 5 point: expert level) [11]. The modified OSATS used in this study [11] excluded the use of assistants and knowledge of instruments and specific procedures from the original OSATS [3]. These items were excluded because the surgeons had sufficient knowledge of the instruments and surgical procedures prior to performing the robotic surgery and did not require a bedside assistant for exposure and additional help.

Videos of RARP were divided into the eight above-mentioned surgical steps, and each part was scored using the video-modified OSATS score. The average of the scores for each part was taken as the total OSATS score for the entire surgery. The three reviewers independently performed scoring, and the average of the three scores was used for further analyses. Surgeons were then divided into two groups: high-OSATS score group (score \geq median) and low-OSATS score group (score $<$ median). In the present study, no measures were taken to instruct the raters or establish standardized rating practices.

Lymph node dissection

Lymph node dissection was performed on patients with $\geq 5\%$ possibility of lymph node metastasis using the Japan prostate cancer nomogram [12]. In most cases, lymph node dissection was performed by a supervisor; therefore, the time spent

on lymph node dissection was excluded from the operation and console times. Similarly, lymph node dissection was excluded from the OSTAS scoring because it was performed by a supervisor.

Data collection

The surgeons' information collected in the present study included their retropubic radical prostatectomy experience, age at the initiation of RARP, and experience with laparoscopic radical prostatectomy. Patient information included patients' age, body mass index (BMI), initial prostate-specific antigen (PSA), Gleason score, prostate weight, and clinical T stage. Information regarding surgical technique, such as means of nerve-sparing, was also obtained. Our dataset did not have missing data.

Endpoints

The primary endpoint was the surgeons achieving the learning curve of the console time. Further, the secondary endpoints were the console time for operating the robot, overall operation time, estimated blood loss, perioperative complication rate, pathological stage, positive surgical margin (PSM) rate, and patients' urinary incontinence at 3 months after surgery were evaluated.

Statistical analysis

The surgical outcomes analyzed were the total operation time, console time, estimated blood loss, prostate volume, complication rate, pathological stage, PSM rate, and urinary incontinence at 3 months post-surgery. Complications were evaluated using the Clavien–Dindo classification system [13]. Outcomes were compared between the high- and low-OSATS score groups using univariate analysis.

To evaluate the learning curve of the console time of RARP, the median and mean console times in the high- and low-OSATS score groups were plotted according to four case series (group 1, 1st–10th cases; group 2, 11th–20th cases; group 3, 21st–30th cases; and group 4, 31st case and beyond), and the two groups were compared.

In addition, cumulative sum analysis (CUSUM) was used to define the learning curves for the respective RARP console times in the high- and low-OSATS groups. This technique plots data from consecutive procedures, transforming raw data into a cumulative sum of differences between the individual values and overall mean. Graphically, this is represented by a curve, with the breakpoint between the ascending and descending portions indicating the number of cases required for the transition from a learning to a proficiency phase [14]. Because the median console value for RARP at

our institution was 129 min, we set the target value for the learning curve at 130 min [10].

Categorical variables were presented as frequency and percentage and continuous variables as median and interquartile range or mean \pm standard deviation. Continuous data were analyzed using the Mann–Whitney U test, and categorical data were analyzed using Pearson's chi-square test and Fisher's exact test. The factors influencing the console time for RARP and continence recovery after RARP were determined using multivariable linear and multivariate logistic regression analyses, respectively.

To assess the reliability of the reviewers' ratings, intra-class correlation coefficients (ICCs) were calculated. Specifically, the ICC of a single random rater and the average ICC for the raters were evaluated.

Statistical analyses were performed using JMP (version 17.0; SAS Institute, Cary, NC, USA) and R software (ver. 4.1.4 R Foundation for Statistical Computing, Vienna, Austria). A P-value < 0.05 was considered statistically significant.

Results

The age of the six surgeons was 27–40 years (mean, 32.3 years) at the start of robotic surgery. They performed 0–40 retropubic radical prostatectomies (mean, 15). None of the surgeons had performed laparoscopic radical prostatectomy prior to the study.

In total, 259 RARP procedures, performed by 6 surgeons, were identified. Among them, one patient who underwent surgery using a retroperitoneal approach was excluded from the analysis. Therefore, we included 43 consecutive RARP procedures for each of the surgeons.

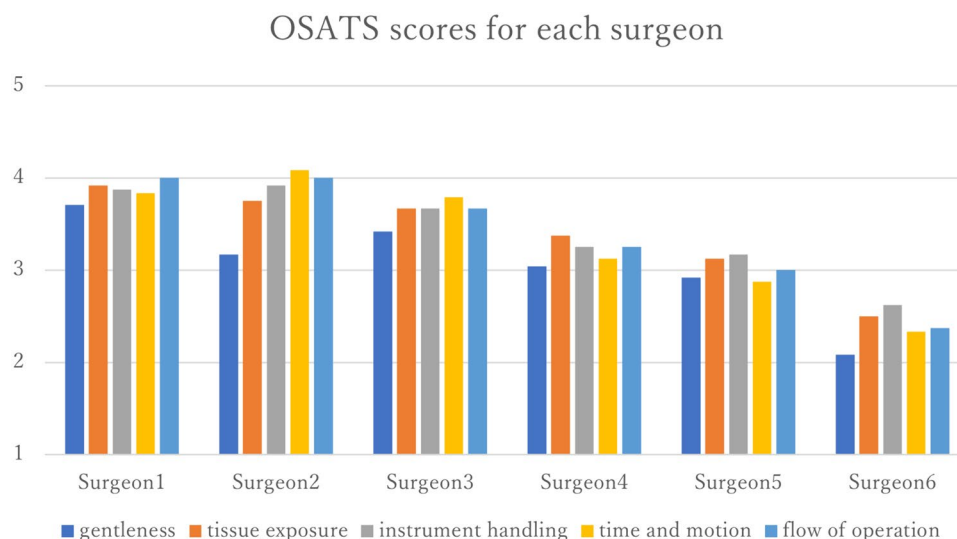
The patient background characteristics of the 10th case of RARP for these surgeons were as follows: mean age, 64.8 years, and mean BMI, 24 kg/m². The mean initial PSA level was 20.5 ng/mL; Gleason score was 6 in one case, 7 in four cases, and 9 in one case. The clinical stage was cT2 in five cases and cT3 in one case; the nerve-sparing technique was used in four cases.

The average surgical time was 213 min, average console time was 152 min, and average blood loss was 240 mL. The pathological stage was pT2 in one case and pT3 in five cases. The average total OSATS score assigned by the three reviewers was 16.6 (11.92–19.33) points.

The details of the OSATS scores for each surgeon are shown in Fig. 1. The ICC for a single random rater was moderate, at 0.45 (95% confidence interval [CI]: 0.36–0.53), whereas the average ICC for the raters showed substantial agreement, at 0.71 (95% CI: 0.63–0.77).

The ranking of the surgeons' scores varied among the three reviewers; however, the members of the two groups

Fig. 1 Average of the modified OSATS score of each surgeon based on three reviewers. OSATS Objective Structured Assessment of Technical Skills



were the same. The scores of all the five components were significantly higher in the high-OSATS score group than in the low-OSATS group (gentleness score: mean 3.4 vs. 2.6 points, tissue exposure score: mean 3.8 vs. 3.0 points, instrument handling score: mean 3.8 vs. 3.0 points, time and motion score: mean 3.9 vs. 2.8 points, and flow of operation score: mean 3.9 vs. 2.9 points; all $P < 0.01$).

On the basis of the OSATS score for each surgeon's 10th RARP case, three surgeons were classified into the high-OSATS score group (18.2–19.3 points) and the remaining three were classified into the low-OSATS score group (11.9–16.0 points). The characteristics of the patients operated by the surgeons categorized into the high- and low-OSATS score groups are presented in Table 1. The initial PSA level and Gleason score were significantly higher for the patients operated by the surgeons in the low-OSATS score group than for those operated by surgeons in the high-OSATS score group ($P = 0.01$ and $P = 0.02$, respectively).

In the univariate analysis, postoperative clinical outcomes showed that the operation and console times were significantly shorter in the high-OSATS score group than in the low-OSATS group (both $P < 0.01$). The continence recovery rate, defined as not wearing pads or using only one safety pad per day, was significantly higher at 3 months for patients operated by the high-OSATS score group surgeons than for those operated by the low-OSATS score group surgeons ($P = 0.03$) (Table 2). However, blood loss, complication rates, and positive margins did not differ significantly between the patients operated by the two groups. Three patients operated by the high-OSATS score surgeons experienced Clavien–Dindo grade 2 complications, including one case of transient fever caused by urinary tract infection and two cases of severe bleeding requiring transfusion. Six patients operated by the low-OSATS score surgeons experienced Clavien–Dindo grade 2 complications,

Table 1 Background demographics of robot-assisted radical prostatectomy cases classified by the surgeon's skill level measured using the modified OSATS

	Modified OSATS score, n (%), or median (IQR)		P-value
	Low, 15.1 (11.9–16.0) points n = 129	High, 18.9 (18.2–19.3) points n = 129	
Age, years	69 (65–72)	69 (64.5–73)	0.85
BMI, kg/m ²	24.3 (22.2–26)	23.4 (22.4–26.1)	0.96
Initial PSA, ng/mL	8.33 (6.3–14.3)	7.0 (5.4–10.2)	0.01
Prostate volume, mL	29.1 (22–38.75)	28 (21.8–32.5)	0.12
Gleason score			0.02
6	6 (5)	13 (10)	
7	62 (48)	76 (59)	
8+	61 (47)	40 (31)	
Clinical stage			0.18
T1c	20 (16)	13 (10)	
T2	101 (78)	112 (87)	
T3	8 (6)	4 (3)	
Nerve-sparing technique			1.00
No	51 (40)	52 (40)	
Yes	78 (60)	77 (60)	

IQR interquartile range, BMI body mass index, PSA prostate-specific antigen, OSATS Objective Structured Assessment of Technical Skills

including four cases of transient fever caused by urinary tract infection and one case each of transient fever due to pneumonia and surgical site infection.

Regarding the console time, multivariable linear regression analyses revealed significant effect in the number of RARPs performed, OSATS score, and prostate

Table 2 Comparison of surgical outcomes after robot-assisted radical prostatectomy between the low- and high-OSATS score groups

	Modified OSATS score, n (%), or median (IQR)		P-value
	Low, 15.1 (11.9–16.0) points n = 129	High, 18.9 (18.2–19.3) points n = 129	
Surgical time, min	202 (175.5–243.5)	158 (138–198.5)	< 0.01
Console time, min	153 (129–190)	112 (98.5–153)	< 0.01
Estimated blood loss, mL	200 (100–400)	200 (100–330)	0.51
Complication total	13 (10.1)	5 (3.9)	0.08
Clavien–Dindo grade I	7 (5.4)	2 (1.6)	0.17
Clavien–Dindo grade II	6 (4.7)	3 (2.3)	0.50
Pathological stage			0.79
pT2	90 (70)	87 (67)	
pT3	39 (30)	42 (33)	
Positive surgical margin			
Total	45 (35)	43 (33)	0.90
pT2	26 (28.9)	19 (21.8)	0.30
pT3	19 (48.7)	24 (57.1)	0.51
Continence status (3 months postoperatively)	38 (29.5)	55 (42.6)	0.03

Continence status was defined as not wearing safety pads or using only one safety pad (20 g) per day
IQR interquartile range, *OSATS* Objective Structured Assessment of Technical Skills

Table 3 Multivariable linear regression analyses of console time

Variable (n = 258)	Regression coefficient (95% CI)	P-value
BMI, kg/m ²	1.22 (–0.43, 2.87)	0.15
Prostate volume, g	0.59 (0.19, 0.99)	< 0.01
Experience of RARP	–2.14 (–2.53, –1.75)	< 0.01
Modified OSATS score (by one point)	–4.51 (–6.39, –2.63)	< 0.01
Nerve sparing technique (non-continuous)	–0.27 (–5.34, 4.80)	0.92

CI confidence interval, *BMI* body mass index, *RARP* robot-assisted radical prostatectomy, *OSATS* Objective Structured Assessment of Technical Skills

volume (regression coefficient [RC] –2.16, $P < 0.01$; RC 16, $P < 0.01$; and RC, 0.38; $P = 0.02$, respectively) (Table 3).

For continence recovery by 3 months after RARP, multivariate regression analyses revealed significant effect in the OSATS score between the two groups (odds ratio [OR]: 0.90 [0.81–1.00], $P = 0.04$) (Table 4).

In the learning curve, the console times for the low- and high-OSATS score groups are shown in Fig. 2. In group 1 (cases 1–10), no difference in the console time was found between the high- and low-OSATS score groups ($P = 0.69$). In group 2 (cases 11–20), group 3 (cases 21–30), and group 4 (cases 31–43), the console time was significantly shorter in the high-OSATS score group than in the low-OSATS score group (all $P < 0.01$).

Analysis of the learning curve on the CUSUM graph showed that the high-OSATS score group reached the breakpoint at 19 cases, whereas the low-OSATS score group did not reach the breakpoint (Fig. 3).

Table 4 Multivariate logistic regression analyses for continence recovery 3 months after RARP

Variable (n = 258)	OR (95% CI)	P-value
Age, years	1.03 (0.98, 1.08)	0.25
BMI, kg/m ²	1.06 (0.97, 1.16)	0.20
Prostate volume, g	1.01 (0.99, 1.03)	0.29
Experience of RARP	1.00 (0.98, 1.03)	0.68
Modified OSATS score (by one point)	0.90 (0.81, 1.00)	0.04
Nerve sparing technique (non-continuous)	1.00 (0.58, 1.70)	0.99

RARP robot-assisted radical prostatectomy, *CI* confidence interval, *OR* odds ratio, *BMI* body mass index, *OSATS* Objective Structured Assessment of Technical Skills

Fig. 2 Comparison of the mean and median console times for the high- and low-OSATS score groups based on the number of experiences among the four groups (the mean scores for each group were plotted). OSATS Objective Structured Assessment of Technical Skills

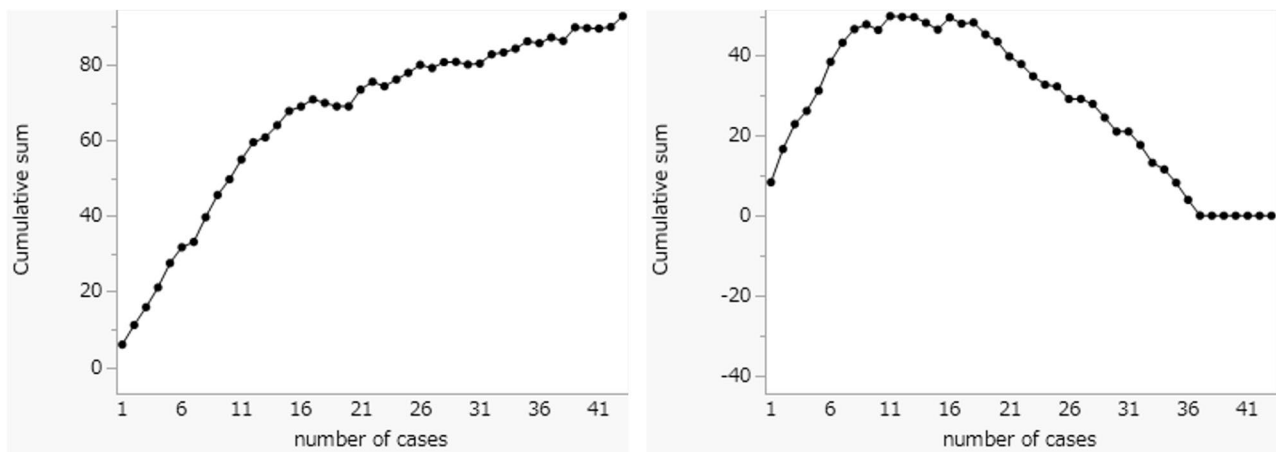
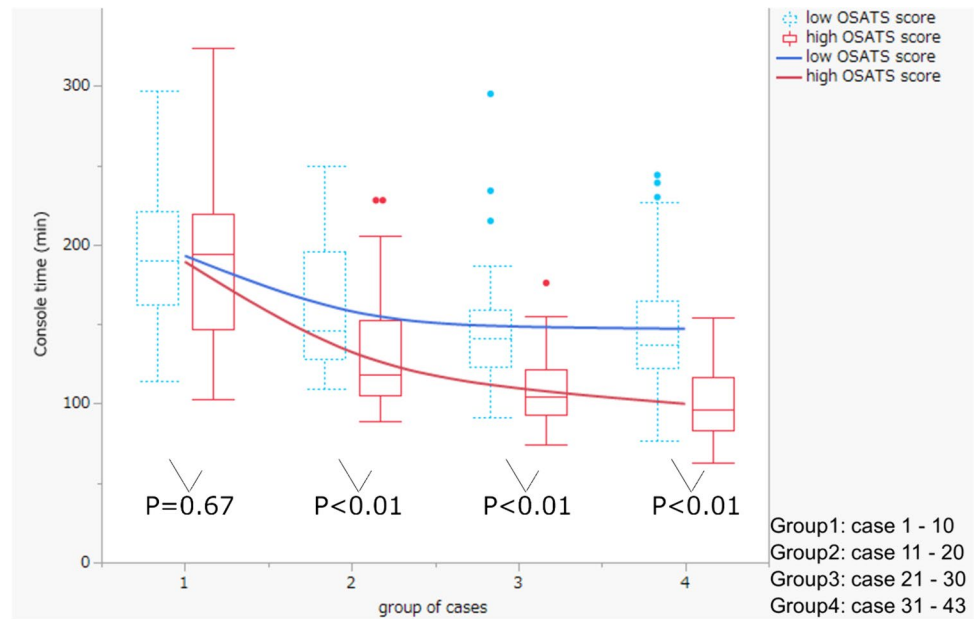


Fig. 3 Cumulative sum analysis curve of the console time for robot-assisted radical prostatectomy according to the number of cases. Right: The learning curve for the high-OSATS score group. The breakpoint was reached at 19 cases. Left: The learning curve for the

low-OSATS score group. The breakpoint of the learning curve was not reached after a total of 43 cases. OSATS Objective Structured Assessment of Technical Skills

Discussion

In this study, we examined the association between the modified OSATS and surgical outcomes of RARP. The group with higher modified OSATS scores for their 10th case of RARP had significantly shorter console and operation times and their patients exhibited earlier continence recovery, compared with the group with lower modified OSATS score. Console time and continence recovery 3 months after surgery were both significantly associated with the modified OSATS score in the multivariate

analysis. The high-OSATS score group also had a significantly shorter console time for the 11th–20th cases than did the low-OSATS score group.

Several video-based surgical assessment methods have been reported [15, 16]. In particular, the Global Assessment Scale (GAS), which focuses on overarching qualities, and procedure-specific operative assessment tool, which separately evaluates key steps and phases of an operation, are both relevant to clinical outcomes. The modified OSATS score used in this study was a type of GAS that has been reported to be associated with clinical outcomes [1, 4, 17]. A previous study evaluated surgeon performance using the

modified OSATS scores for laparoscopic bariatric surgery [11]. They found differences in the complication rate (14.5% vs. 5.2%, $P < 0.001$), mortality (0.26% vs. 0.05%, $P = 0.01$), operation time (137 min vs. 98 min, $P < 0.001$), reoperation rates (3.4% vs. 1.6%, $P = 0.01$), and readmission rates (6.3% vs. 2.7%, $P < 0.001$) between the 25% of surgeons with the highest modified OSATS ratings and 25% of surgeons with the lowest ratings [11]. Associations between surgical performance and clinical outcomes have been reported not only in laparoscopic surgery but also in robot-assisted surgery [18, 19]. Specifically, the evaluation of intraoperative videos using the OSATS scores for robotic pancreatoduodenectomy revealed a relationship with clinical outcomes. When 153 cases of robotic pancreatoduodenectomy were evaluated using the OSATS scores, the incidence of postoperative pancreatic fistula decreased for patients operated by surgeons with higher OSATS scores (OR: 4.01, $P = 0.004$) [19].

In our study of six surgeons, those with higher OSATS scores had shorter console and operation times and their patients had better urinary continence by 3 months after surgery, compared with the surgeons with lower OSATS scores. Our results are consistent with those of previous studies. Although previous studies have evaluated a video of a “typical” case for each surgeon, we evaluated the OSATS score for the surgeons’ 10th case of RARP [11, 20, 21].

A recent review indicated that reaching the learning curve of the console time requires 16–300 cases [22]. Our study, which evaluated the surgical technique of surgeons accustomed to robotic surgery, revealed that the surgeon’s initial technique affected their learning curve and patients’ clinical outcomes. In addition, the high-OSATS score group reached the breakpoint after 19 cases, whereas the low-OSATS score group did not reach the breakpoint after 43 cases. The high-OSATS group showed results similar to those of previous studies [23]. The group with a higher OSATS score for the 10th case had a shorter console time for the 11th–20th cases and beyond, compared with the group with a lower OSATS score for the 10th case. Thus, subsequent learning curves can be predicted on the basis of earlier surgical skills.

In other studies, surgeons from multiple centers were selected, which may have employed different teaching methods and surgical procedure details [11]. In our study, we were able to evaluate the effect of the surgeon’s skill because we selected surgeons who had received the same education and used the same surgical procedures [7].

Evaluation of surgical skills is required to improve surgical education [24]. As each individual has a different learning curve, predicting a surgeon’s learning curve by evaluating their surgical skills may help to identify surgeons who require more surgical education and training [25, 26]. Various methods have been reported for education in RARP, including training programs, simulation training using wet lab models or phantoms, and virtual reality [27–31]. Clinical

outcomes have also been reported to improve with training in RARP, and thus surgeons with low initial OSATS scores may need to use these training tools to improve their skills [27].

Moreover, the relationship between surgical performance and early continence recovery after RARP has been reported. When surgery was evaluated using the Global Evaluative Assessment of Robotic Skill (GEARS), a tool developed to evaluate robotic surgery skills, the GEARS scores for the bladder neck and vesicourethral anastomosis procedures were particularly related to early continence recovery at 3 months postoperatively (OR 0.69, 95% CI 0.51–0.94, OR 0.70, 95% CI 0.50–0.97, respectively) [32, 33]. Although their evaluation method differed from that used in our study, the results indicated that good surgical performance may be related to early continence recovery postoperatively, which is consistent with our results [33]. In contrast, we found no association of nerve preservation, prostate size, or the number of previous surgeries with urinary continence. It is possible that the breakpoint of the learning curve for nerve preservation had not been reached or that the number of cases was insufficient to achieve it.

The GEARS includes the items “depth perception,” “bimanual dexterity,” “efficiency,” “force sensitivity,” “autonomy,” and “robotic control” [32], whereas the modified OSATS comprises the items “gentleness,” “tissue exposure,” “instrument handling,” and “time.” Although there is no report on the merits of either the GEARS or OSATS, we thought that the modified OSATS was more appropriate for evaluating the surgical skills of novice surgeons, based on its constituent items. Nevertheless, as the OSATS has been reported to be correlated to the GEARS, either tool could provide valid results [34].

This study has some limitations. First, the reviewers had not received prior training, which caused some variation in the ICC. Nevertheless, we believe the validity of the results was maintained because all three reviewers agreed on the composition of the high- and low-OSATS score groups. Second, no effect of the OSATS score of the surgeons on the blood loss, complications, and positive surgical margins in the patients were found in the present study. A recent review showed that it requires more experience to reach the breakpoint of the learning curve for bleeding, complications, and positive surgical margins, compared with operative time [22]. It is possible that the number of cases in this study was too low to observe statistical significance in this respect. Finally, owing to the small number of surgeons who met the criteria, only six surgeons could be evaluated. Further research with an increased number of surgeons and surgical cases is required.

In conclusion, based on the study results, the initial technical skills of surgeons may be used to predict their learning curve in RARP and the patients’ clinical outcomes.

Acknowledgements None.

Author contributions Conception and design of the study: M Yamazaki, H Kawahira, K Oiwa, Y Maeda. Data acquisition: M Yamazaki. Data analysis: M Yamazaki, H Yokoyama, T Kameda. Data interpretation: M Yamazaki, H Kawahira, K Oiwa, Y Maeda. Drafting of the article: M Yamazaki. Critical revisions of the manuscript: J Kamei, T Sugihara, S Ando, T Fujimura. All authors have read and approved the final version of the paper for publication.

Funding None.

Declarations

Disclosures Masahiro Yamazaki, Hiroshi Kawahira, Yoshitaka Maeda, Kosuke Oiwa, Hirotaka Yokoyama, Tomohiro Kameda, Jun Kamei, Toru Sugihara, Satoshi Ando, and Tetsuya Fujimura have no conflicts of interest to disclose.

References

1. Varban OA, Thumma JR, Finks JF, Carlin AM, Ghaferi AA, Dimick JB (2021) Evaluating the effect of surgical skill on outcomes for laparoscopic sleeve gastrectomy: a video-based study. *Ann Surg* 273:766–771. <https://doi.org/10.1097/sla.00000000000003385>
2. Fecso AB, Bhatti JA, Stotland PK, Quereshy FA, Grantcharov TP (2019) Technical performance as a predictor of clinical outcomes in laparoscopic gastric cancer surgery. *Ann Surg* 270:115–120. <https://doi.org/10.1097/sla.0000000000002741>
3. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 84:273–278. <https://doi.org/10.1046/j.1365-2168.1997.02502.x>
4. Grüter AAJ, Van Lieshout AS, van Oostendorp SE, Henckens SPG, Ket JCF, Gisbertz SS, Toorenvliet BR, Tanis PJ, Bonjer HJ, Tuynman JB (2023) Video-based tools for surgical quality assessment of technical skills in laparoscopic procedures: a systematic review. *Surg Endosc* 37:4279–4297. <https://doi.org/10.1007/s00464-023-10076-z>
5. Boal MWE, Anastasiou D, Tesfai F, Ghamrawi W, Mazomenos E, Curtis N, Collins JW, Sridhar A, Kelly J, Stoyanov D, Francis NK (2024) Evaluation of objective tools and artificial intelligence in robotic surgery technical skills assessment: a systematic review. *Br J Surg* 111:znad331. <https://doi.org/10.1093/bjs/znad331>
6. Nguyen-Nielsen M, Borre M (2016) Diagnostic and therapeutic strategies for prostate cancer. *Semin Nucl Med* 46:484–490. <https://doi.org/10.1053/j.semnuclmed.2016.07.002>
7. Ahlring TE, Skarecky D, Lee D, Clayman RV (2003) Successful transfer of open surgical skills to a laparoscopic environment using a robotic interface: initial experience with laparoscopic radical prostatectomy. *J Urol* 170:1738–1741. <https://doi.org/10.1097/01.ju.0000092881.24608.5e>
8. Shrivastava A, Baliga M, Menon M (2007) The Vattikuti Institute prostatectomy. *BJU Int* 99:1173–1189. <https://doi.org/10.1111/j.1464-410X.2007.06878.x>
9. Fujimura T, Menon M, Fukuhara H, Kume H, Suzuki M, Yamada Y, Niimi A, Nakagawa T, Igawa Y, Homma Y (2016) Validation of an educational program balancing surgeon training and surgical quality control during robot-assisted radical prostatectomy. *Int J Urol* 23:160–166. <https://doi.org/10.1111/iju.12993>
10. Ando S, Sugihara T, Hinotsu S, Kishino H, Hirata D, Watanabe R, Yanase A, Yokoyama H, Hoshina H, Endo K, Kamei J, Takaoka E, Fujimura T (2024) Early recovery of urinary continence after robot-assisted radical prostatectomy is associated with membranous urethra and neurovascular bundle preservation. *Int J Urol*. <https://doi.org/10.1111/iju.15388>
11. Naito S, Kuroiwa K, Kinukawa N, Goto K, Koga H, Ogawa O, Murai M, Shiraiishi T (2008) Validation of Partin tables and development of a preoperative nomogram for Japanese patients with clinically localized prostate cancer using 2005 International Society of Urological Pathology consensus on Gleason grading: data from the Clinicopathological Research Group for Localized Prostate Cancer. *J Urol* 180:904–909. <https://doi.org/10.1016/j.juro.2008.05.047>. (discussion 909–910)
12. Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR, Dimick J, Banerjee M, Birkmeyer NJ (2013) Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 369:1434–1442. <https://doi.org/10.1056/NEJMsa1300625>
13. Dindo D, Demartines N, Clavien PA (2004) Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 240:205–213. <https://doi.org/10.1097/01.sla.0000133083.54934.ae>
14. David JB, Mathieu RR, Gaëlle GP, Rémy SN, Raphaël P (2007) Quality control of surgical and interventional procedures: a review of the CUSUM. *Qual Saf Health Care* 16:203–207. <https://doi.org/10.1136/qshc.2006.020776>
15. Nathan JC, Jake DF, Danilo M, Chris SBB, Peter JH, Sarah A, George BH, Andrew RLS, Nader KF (2020) Association of surgical skill assessment with clinical outcomes in cancer surgery. *JAMA Surg* 155:590–598. <https://doi.org/10.1001/jamasurg.2020.1004>
16. Heinrich H, Lindsay S, Eliane MS, Guylaine GL, Teodor PG (2015) The generic error rating tool: a novel approach to assessment of performance and surgical education in gynecologic laparoscopy. *J Surg Educ* 72:1259–1265. <https://doi.org/10.1016/j.jsurg.2015.04.029>
17. Kurashima Y, Kitagami H, Teramura K, Poudel S, Ebihara Y, Inaki N, Nakamura F, Misawa K, Shibao K, Nagai E, Akashi Y, Takiguchi S, Kobayashi S, Hirano S (2022) Validation study of a skill assessment tool for education and outcome prediction of laparoscopic distal gastrectomy. *Surg Endosc* 36:8807–8816. <https://doi.org/10.1007/s00464-022-09305-8>
18. Kaoukabani G, Gokcal F, Fanta A, Liu X, Shields M, Stricklin C, Friedman A, Kudsi OY (2023) A multifactorial evaluation of objective performance indicators and video analysis in the context of case complexity and clinical outcomes in robotic-assisted cholecystectomy. *Surg Endosc* 37:8540–8551. <https://doi.org/10.1007/s00464-023-10432-z>
19. van den Broek BLJ, Zwart MJW, Bonsing BA, Busch OR, van Dam JL, de Hingh I, Hogg ME, Luyer MD, Mieog JSD, Stibbe LA, Takagi K, Tran TCK, de Wilde RF, Zeh HJ 3rd, Zureikat AH, Groot Koerkamp B, Besselink MG (2023) Video grading of pancreatic anastomoses during robotic pancreatoduodenectomy to assess both learning curve and the risk of pancreatic fistula: a post hoc analysis of the LAELAPS-3 training program. *Ann Surg* 278:e1048–e1054. <https://doi.org/10.1097/sla.0000000000005796>
20. Scally CP, Varban OA, Carlin AM, Birkmeyer JD, Dimick JB (2016) Video ratings of surgical skill and late outcomes of bariatric surgery. *JAMA Surg* 151:e160428. <https://doi.org/10.1001/jamasurg.2016.0428>
21. Varban OA, Thumma JR, Carlin AM, Finks JF, Ghaferi AA, Dimick JB (2020) Peer assessment of operative videos with sleeve gastrectomy to determine optimal operative technique. *J A Coll Surg* 231:470–477. <https://doi.org/10.1016/j.jamcollsurg.2020.06.016>
22. Grivas N, Zachos I, Georgiadis G, Karavitakis M, Tzortzis V, Mamoulakis C (2022) Learning curves in laparoscopic

- and robot-assisted prostate surgery: a systematic search and review. *World J Urol* 40:929–949. <https://doi.org/10.1007/s00345-021-03815-1>
23. Catherine L, Giacomo N, Alex M, Khurshid AG, Matthew B, Ben C, Richard P, Johar R, Henk VP, James P, Prokar D, Kamran A (2016) Structured and modular training pathway for robot-assisted radical prostatectomy (RARP): validation of the RARP assessment score and learning curve assessment. *Eur Urol* 69:526–535. <https://doi.org/10.1016/j.eururo.2015.10.048>
 24. Matsuda T, Yoshida K, Habuchi T, Kinoshita H, Kanayama HO, Terachi T (2017) Scientific skill assessment of basic surgical dissection and overall laparoscopic performance. *J Endourol* 31:S25–S29. <https://doi.org/10.1089/end.2016.0554>
 25. Gandi C, Totaro A, Bientinesi R, Marino F, Pierconti F, Martini M, Russo A, Racioppi M, Bassi P, Sacco E (2022) A multi-surgeon learning curve analysis of overall and site-specific positive surgical margins after RARP and implications for training. *J Robot Surg* 16:1451–1461. <https://doi.org/10.1007/s11701-022-01378-w>
 26. Talme M, Harilainen A, Lindahl J, Bister V (2023) Anterior cruciate ligament reconstruction learning curve study—comparison of the first 50 consecutive patients of five orthopaedic surgeons during a 5-year follow-up. *Knee* 44:1–10. <https://doi.org/10.1016/j.knee.2023.06.011>
 27. Wang F, Zhang C, Guo F, Sheng X, Ji J, Xu Y, Cao Z, Lyu J, Lu X, Yang B (2021) The application of virtual reality training for anastomosis during robot-assisted radical prostatectomy. *Asian J Urol* 8:204–208. <https://doi.org/10.1016/j.ajur.2019.11.005>
 28. Lumen N, Van Praet C, De Troyer B, Fonteyne V, Oosterlinck W, Decaestecker K, Mottrie A (2013) Safe introduction of robot-assisted radical prostatectomy after a training program in a high-volume robotic centre. *Urol Int* 91:145–152. <https://doi.org/10.1159/000350652>
 29. Cacciamani G, De Marco V, Siracusano S, De Marchi D, Bizzotto L, Cerruto MA, Motton G, Porcaro AB, Artibani W (2017) A new training model for robot-assisted urethrovesical anastomosis and posterior muscle-fascial reconstruction: the Verona training technique. *J Robot Surg* 11:123–128. <https://doi.org/10.1007/s11701-016-0626-4>
 30. Witthaus MW, Farooq S, Melnyk R, Campbell T, Saba P, Mathews E, Ezzat B, Ertefaie A, Frye TP, Wu G, Rashid H, Joseph JV, Ghazi A (2020) Incorporation and validation of clinically relevant performance metrics of simulation (CRPMS) into a novel full-immersion simulation platform for nerve-sparing robot-assisted radical prostatectomy (NS-RARP) utilizing three-dimensional printing and hydrogel casting technology. *BJU Int* 125:322–332. <https://doi.org/10.1111/bju.14940>
 31. Johnson BA, Timberlake M, Steinberg RL, Kosemund M, Mueller B, Gahan JC (2019) Design and validation of a low-cost, high-fidelity model for urethrovesical anastomosis in radical prostatectomy. *J Endourol* 33:331–336. <https://doi.org/10.1089/end.2018.0871>
 32. Alvin CG, David WG, James CS, Brian JM, Brian JD (2012) Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. *J Urol* 187:247–252. <https://doi.org/10.1016/j.juro.2011.09.032>
 33. Goldenberg MG, Goldenberg L, Grantcharov TP (2017) Surgeon performance predicts early continence after robot-assisted radical prostatectomy. *J Endourol* 31:858–863. <https://doi.org/10.1089/end.2017.0284>
 34. Hyeong WY, Jin WY, Chan YS, Jong-kyu K, In EB, Hyungju K, Young JC, Su-jin K, June YC, Kyu EL (2018) Development of a surgical training model for bilateral axillobreast approach robotic thyroidectomy. *Surg Endosc* 32:1360–1367. <https://doi.org/10.1007/s00464-017-5816-2>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.