


Prior experience in laparoscopic rectal surgery can minimise the learning curve for robotic rectal resections: a cumulative sum analysis

Manfred Odermatt¹  · Jamil Ahmed² · Sofoklis Panteleimonitis² · Jim Khan¹ · Amjad Parvaiz^{3,4}

Received: 15 September 2016 / Accepted: 3 February 2017 / Published online: 7 March 2017
© Springer Science+Business Media New York 2017

Abstract

Background The learning curve for robotic colorectal surgery is ill-defined. This study aimed to investigate the learning curve of experienced laparoscopic rectal surgeons when starting with robotic total mesorectal excision (TME) using cumulative sum (CUSUM) charts.

Methods This retrospective case series analysed patients who underwent curative and elective laparoscopic or robotic TMEs for rectal cancer performed by two surgeons. The first consecutive robotic TME cases of each surgeon were 1:1 propensity score matched to their laparoscopic TME cases using age, body mass index, American Society of Anesthesiologists grade, T stage (AJCC) and tumour location height. The matched laparoscopic cases defined individual standards for the quality indicators: operating time, R stage, lymph node harvest, length of hospital stay and major complications (Clavien–Dindo grade 3–5). Deviation of more than a quarter of a standard deviation from the mean for the continuous indicators, or exceeding the observed risk for the binary indicators was defined as off-target with an upward inflection in the CUSUM curve.

Results From 2006 to 2015, 384 (294 laparoscopic; 90 robotic) TMEs met the inclusion criteria. Surgeon A performed 206 (70.1%) of the laparoscopic and 43 (47.8%) of the robotic cases. Surgeon B performed 88 (29.9%) of the laparoscopic and 47 (52.2%) of the robotic cases. After matching, no covariate exhibited an absolute standardised mean difference >0.25. For surgeon A, the CUSUM curves showed no apparent learning process compared to his laparoscopic standards. For surgeon B, a learning process for operation time, lymph node harvest and major complications was demonstrated by an initial upward inflection of the CUSUM curves; after 15 cases, all quality indicators were generally on target.

Conclusions For experienced laparoscopic colorectal surgeons, the formal learning process for robotic TME may be short to reach a similar performance level as obtained in conventional laparoscopy.

Keywords Robotic surgery · Total mesorectal excision · Rectal surgery · Cumulative sum · Learning curve

✉ Manfred Odermatt
manfred.odermatt@gmx.ch

¹ Minimally Invasive Colorectal Unit, Queen Alexandra Hospital, Southwick Hill Road, Cosham, Portsmouth PO6 3LY, UK

² Poole Hospital NHS Foundation Trust, Poole, UK

³ FRCS (Gen) FRCS, European Academy of Robotic Colorectal Surgery, Poole Hospital NHS Foundation Trust, Poole, UK

⁴ Laparoscopic and Robotic Colorectal Surgery Champalimaud Foundation, Lisbon, Portugal

The robotic system is used in various fields of surgery and its application to different indications continues to expand in parallel with the development of technology [1]. Feasibility, safety and auspicious outcomes of robotic total mesorectal excision (TME) have been demonstrated by various case series [2–7]. Preliminary data from the ROLARR trial (Robotic versus Laparoscopic Resection for Rectal Cancer) has shown that robotic TME does not increase the conversion rate, though no statistically significant clinical or oncological benefit has been reported so far [8]. Before taking part in the ROLARR trial, surgeons had performed an average of 91 laparoscopic versus only 25 robotic resections which might suggest that some of the participating

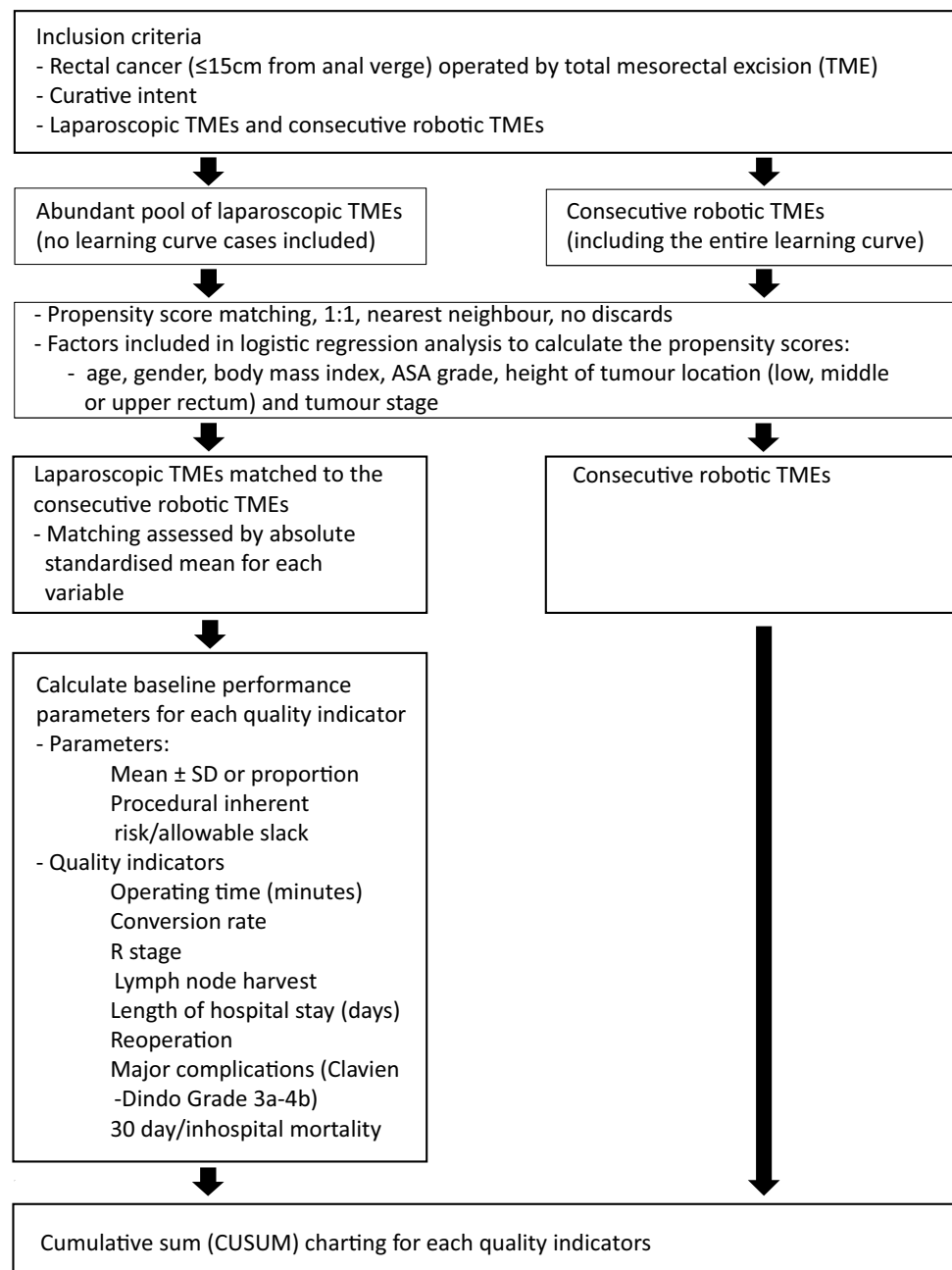
surgeons were still in their learning curve for robotic surgery whilst already being an expert in laparoscopic rectal surgery during this trial. In fact, there is only sparse data available about the robotic learning curve of expert laparoscopic rectal surgeons which might be different to that for a rectal surgeon primarily starting with robotic TME without prior laparoscopic experience [9]. However, this situation is certainly less likely as in most surgical centres only experienced laparoscopic colorectal surgeons adopt the robotic technique. Hence, the learning curve of an expert laparoscopic TME surgeon who starts robotic TME is likely of paramount interest for colorectal services, especially as the

robotic system theoretically should simplify the operative procedure which may result in a favourable learning curve. The scope of this study is to produce illustrative robotic learning curves for two experienced laparoscopic colorectal surgeons using cumulative sum (CUSUM) charts, whilst adjusting for confounders.

Materials and methods

A retrospective analysis of a prospectively maintained database was carried out (Fig. 1). Laparoscopic and

Fig. 1 Flow chart of the method used to assess the learning process of an individual surgeon



robotic total mesorectal resections (anterior resections) for non-metastatic rectal cancers performed from October 2006 to November 2015 at the Minimally Invasive Colorectal Unit, Queen Alexandra Hospital, Portsmouth, were included in this study. Approximately, 250–300 colorectal resections per annum are performed in our unit. During the study period, a dual console four arm robotic system (Da Vinci® Si, Intuitive Surgical®) was introduced and two experienced laparoscopic colorectal surgeons (surgeon A and B) who were also trainers for the National Training Programme (NTP) for Laparoscopic Colorectal Surgery in England (Lapco) adopted the robotic method. Both surgeons had reached a plateau in the learning curve with respect to their independent and proficient performance of a laparoscopic colorectal resection, as to Lapco standards (<http://www.lapco.nhs.uk>). The laparoscopic experience of surgeon A and B before starting robotics consisted of approximately 1500 and 400 colorectal procedures, respectively. The robotic cases for each individual surgeon were consecutive and therefore included the complete learning curve. A full-time research assistant collected all variables of interest in a structured database and validated the data, including ensuring consistency with patient notes and checking death records for patients lost to follow-up. For each individual surgeon, the robotic cases were 1:1 nearest neighbour propensity score matched to his laparoscopic cases to obtain comparable groups of patients, with the laparoscopic cases serving as a reference to calculate the proficiency targets for the robotic cases. The variables used to calculate the propensity scores were age, gender, body mass index (BMI), American Society of Anesthesiologists (ASA) grade, height of tumour location (low, middle or upper rectum) and American Joint Committee on Cancer (AJCC) tumour stage. Outcome variables of interest, defining quality indicators, were operation time (minutes), conversion rate, lymph node harvest (count), length of hospital stay (days), reoperation rate, the presence or absence of major complications (defined as Clavien–Dindo Grade 3a–4b) and 30-day/inhospital mortality (Clavien–Dindo grade 5) [10]. For each surgeon, the mean and standard deviations of the continuous quality indicators were calculated for his laparoscopic cases to define the individual baseline proficiency against which the robotic cases were compared when performing the cumulative sum analysis (CUSUM). For the binary variables, the rate of each outcome for the laparoscopic cases was used as the baseline/inherent risk, again calculated separately for each surgeon. Potential sources of bias were addressed by focusing on treatment of non-metastatic rectal cancer with curative intent i.e. anterior resection with TME and by the propensity score matching providing a comparable laparoscopic group with similar operative and perioperative risk factors to avoid selection bias (due to case mix). The study size

evolved from the number of consecutive robotic cases of the involved surgeons matched to an equivalent number of non-learning curve laparoscopic cases. Quantitative variables were analysed as such with the exception of tumour height from anal verge which was categorised as low (2–5 cm), medium (6–11 cm) or high (12–15 cm) rectal cancers. Operating time was defined as the time from incision of the skin until the final stitch which also included the docking time in robotic cases. All included patients signed an informed consent allowing their data to be used for retrospective analysis and its publishing. The requirements for anonymization of personal dataset by the Data Protection Act 1998 were satisfied. According to the Health Research Authority (HRA), this study was not classified to need their approval as it is an audit.

Operative procedure

The laparoscopic TMEs were performed following a standardised approach. This included medial to lateral mobilisation, high tie of the inferior mesenteric artery, division of the inferior mesenteric vein at the lower border of the pancreas, complete mobilisation of the splenic flexure in all cases and total mesorectal excision using sharp diathermy dissection usually followed by a colorectal stapler anastomosis. The robotic procedures were performed following similar principles, again using a highly standardised approach including complete mobilisation of the splenic flexure in all cases [11]. A single-dock method was applied, and an experienced robotic surgeon proctored the first ten robotic TMEs of each surgeon. Although a dual console was used for proctoring, it was not intended for the proctor to take over parts of the operation, but only to demonstrate technical solutions and ensure a safe operation. Observations and simulator training (of at least 30 h) were carried out beforehand. Both surgeon A and B had colorectal fellows who assisted with the robotic cases. They were different but the team for each surgeon was generally consistent. Although the assistants had no previous robotic experience, they had extensive previous exposure to advanced laparoscopic colorectal surgery.

Cumulative sum charting

Cumulative sum (CUSUM) curves for each quality indicator of interest were used to monitor the performance of each surgeon [12, 13]; the presentation format described by Kestin was used [14]. The x-axis represents the consecutive case number and the y-axis represents the CUSUM score. For the binary outcomes (X), the CUSUM score is the cumulative sum of $X_i - X_0$ where X_i represents the success ($X_i = 0$) or failure ($X_i = 1$) of each consecutive procedure. X_0 is the inherent risk of failure of the procedure, which in

our study has been calculated using each surgeon's matched laparoscopic reference group as a baseline (Fig. 1). For the continuous quality indicators (T), such as operating time, the CUSUM score is the cumulative sum of $T_i - (x + k)$, where T_i is the outcome of each successive procedure; k is the tolerable slack from the reference mean (\bar{x}) in the laparoscopic group and was defined as a quarter of a standard deviation in order to set tight targets and to account for the dispersion of the reference data. A tolerable shift of half a standard deviation is a common rule-of-thumb in industrial processes where, however, the standard deviation is normally derived from a larger sample size [15]. The CUSUM curves ascend when the set targets are not reached which reflects an ongoing learning process; the curve runs more or less parallel to the x-axis when the performance is similar to the laparoscopic standards, reflecting no learning process; the curve has a downward trend when the performance is more often on target than expected.

Statistical methods

Propensity scores were calculated via logistic regression analysis with method of procedure (laparoscopic versus robotic) as the outcome variable and then nearest neighbour 1:1 matching applied (with no discards, nor a specified caliper distance). The level of balance achieved was assessed by the absolute standardised mean difference for each baseline variable, the multivariate imbalance measure L1 (Iacus, King and Porro, 2010) and the overall balance test (Hansen and Bowers, 2010). For the matching process, PS Match for SPSS (Version 1.0, by Felix Thoemmes, 2011) [16] was used; underlying R packages included MatchIt, Rltools, and cem [17–20]. For the outcomes, categorical variables were compared using a Pearson Chi-square or Fisher's exact test, as appropriate, and continuous variables were compared using the non-parametric, Mann–Whitney U test. The statistical analyses were performed using SPSS (Statistical Package for Social Sciences, Version 20, IBM Corp. Released 2011).

Results

A total of 384 (294 laparoscopic and 90 robotic) total mesorectal resections met the inclusion criteria. Surgeon A performed 206 (70.1%) of the laparoscopic and 43 (47.8%) of the robotic cases, whilst surgeon B performed 88 (29.9%) of the laparoscopic and 47 (52.2%) of the robotic cases during the study period. After propensity score matching, no baseline variable exhibited an absolute standardised mean difference >0.25 ($|d| > 0.25$), either for surgeon A or surgeon B. The multivariate imbalance measure L1 showed an increase in balance (0.976 and 0.955 before and 0.837

and 0.915 after matching for surgeon A and B, respectively), and the overall balance test showed no significant difference after matching ($p=0.963$ and $p=0.988$, respectively). Biometric, oncologic and operative baseline characteristics of the robotic and matched laparoscopic cases of each surgeon were similar and are summarised in Table 1. Overall, surgeon A treated more low rectal cancers (47 vs. 32%) but less cases with neoadjuvant radiotherapy (17 vs. 30%) when compared to surgeon B. The outcomes of interest, including the quality indicators, are summarised in Table 2. On an individual level, there was no significant difference with respect to operation time, R stage, lymph node harvest, length of hospital stay, conversion rate and major complications between the laparoscopic and robotic cases for both surgeons. Only the readmission rate was statistically significantly higher in the laparoscopic compared with the robotic group for surgeon B (21 vs. 6%, $p=0.035$). The cut-off values for robotic on-target performance for the continuous quality indicators, and the expected inherent risk for the binary quality indicators, were calculated based upon the laparoscopic outcomes and are listed for each quality indicator in Table 3. The on-target cut-off for operation time was considerably higher for surgeon B than for surgeon A (292 vs. 233 min) as was the inherent risk for major complications Clavien–Dindo 3b–5 (15 vs. 7%); on the other hand, surgeon B had a slightly higher lymph node harvest cut-off (15 versus 14) and also a shorter hospital stay cut-off (11 vs. 12 days). Finally, the CUSUM charts for the quality indicators are shown in Fig. 2. As the conversion rate was zero in the robotic group for both surgeons and only one conversion occurred in the laparoscopic group, no CUSUM chart was drawn for this quality indicator. For similar reasons (i.e. due to small numbers of events), mortality was included in the major complications outcome so that this constituted a single quality indicator (Clavien–Dindo grade 3b–5).

For surgeon A who had performed more laparoscopic TMEs (1500 laparoscopic colorectal procedures before starting robotics and 206 cases meeting the inclusion criteria during the study) before starting robotic TME, the CUSUM curves showed only a short learning process for operation time, necessitating seven tutored cases, whilst there was apparently minimal to no learning process for the other quality indicators, such as lymph node harvest, length of stay and major complications. In fact, the general downward trend for operation time, lymph node harvest and length of stay indicates that these quality indicators were more often on rather than off their pre-defined targets.

For surgeon B who had performed less laparoscopic TMEs beforehand (400 laparoscopic colorectal procedures before starting robotics and 88 cases meeting the inclusion criteria during the study), the CUSUM curves showed a clear learning process for operation time, length of stay and

Table 1 Baseline characteristics of patients who underwent laparoscopic or robotic total mesorectal resection by two surgeons after propensity score matching

	Surgeon A		Surgeon B	
	Laparoscopic (n=43)	Robotic (n=43)	Laparoscopic (n=47)	Robotic (n=47)
Age [median (min–max)] in years	68 (43–86)	67 (53–92)	67 (44–86)	68 (35–88)
Gender				
Male	29 (67.4%)	30 (69.8%)	35 (74.5%)	34 (72.3%)
Female	14 (32.6%)	13 (30.2%)	12 (25.5%)	13 (27.7%)
Body mass index [median (min–max)] in kg/m ²	26 (20–43)	27 (21–46)	27 (19–43)	27 (20–41)
ASA ^a grade				
1	6 (14.0%)	4 (9.3%)	12 (25.5%)	7 (14.9%)
2	31 (72.1%)	33 (76.7%)	27 (57.4%)	35 (74.5%)
3	6 (14.0%)	6 (14.0%)	8 (17.0%)	5 (10.6%)
Height of tumour location				
0–6 cm (low)	20 (46.5%)	20 (46.5%)	16 (34.0%)	14 (29.8%)
7–11 cm (middle)	18 (41.9%)	19 (44.2%)	20 (42.6%)	25 (53.2%)
12–15 cm (upper)	5 (11.6%)	4 (9.3%)	11 (23.4%)	8 (17.2%)
T stage				
T0	0 (0.0%)	1 (2.3%)	3 (6.4%)	4 (8.5%)
T1	4 (9.3%)	5 (11.6%)	6 (12.8%)	8 (17.0%)
T2	13 (30.2%)	15 (34.9%)	10 (21.3%)	17 (36.2%)
T3	24 (55.8%)	20 (46.5%)	27 (57.4%)	17 (36.2%)
T4	2 (4.7%)	2 (4.7%)	1 (2.1%)	1 (2.1%)
Tumour stage				
Stage 0	0 (0.0%)	1 (2.3%)	3 (6.4%)	4 (8.5%)
Stage 1	10 (23.3%)	13 (30.2%)	15 (31.9%)	20 (42.6%)
Stage 2	13 (30.2%)	12 (27.9%)	17 (36.2%)	7 (14.9%)
Stage 3	20 (46.5%)	17 (39.5%)	12 (25.5%)	16 (34.0%)
Neoadjuvant radiotherapy				
None	37 (86.0%)	34 (79.1%)	33 (70.2%)	33 (70.2%)
Short course	0 (0.0%)	2 (4.7%)	2 (4.3%)	1 (2.1%)
Long course	6 (14.0%)	7 (16.3%)	12 (25.5%)	13 (27.7%)
Stoma				
None	1 (2.3%)	2 (4.7%)	6 (12.8%)	5 (10.6%)
Ileostomy	40 (93.0%)	38 (88.4%)	41 (87.2%)	40 (85.1%)
Colostomy	2 (4.6%)	3 (7.0%)	0 (0.0%)	2 (4.3%)

^aAmerican Society of Anesthesiologists

major complications. However, all of these indicators were ultimately on-target, showing a comparable performance to laparoscopy after just 15 robotic procedures. No evidence of a learning process was observed for lymph node harvest. Other systematic influences explaining the learning curve like a sudden change of the assistant could not be found.

Discussion

This study suggests that experienced laparoscopic colorectal surgeons may undergo a short learning process when

changing from laparoscopic to robotic TME. Furthermore, the number of previous laparoscopic TMEs performed may also influence the number of cases needed to reach an equivalent performance level in robotic TME. This is well demonstrated by the CUSUM curves running parallel to the x-axis or inflecting downwards after a maximum of only 15 robotic cases, whilst the more experienced surgeon (A) had a shorter learning process than the less-experienced surgeon (B). The results also show that the introduction of a robotic system into a specialist colorectal unit may only have some marginal effect on case load per time period and short-term outcomes. In case of extensive laparoscopic

Table 2 Short-term outcomes of consecutive learning curve robotic total mesorectal excisions compared to non-learning curve conventional laparoscopic resections performed by surgeon A and B

	Surgeon A			Surgeon B		
	Laparoscopic (n=43)	Robotic (n=43)	p value	Laparoscopic (n=47)	Robotic (n=47)	p value
Operation time [median (range)] in minutes	220 (145–420)	210 (140–340)	0.198 ^a	290 (160–420)	280 (170–456)	0.587 ^a
R positive	1 (2.3%)	1 (2.3%)	1.000 ^b	4 (8.5%)	1 (2.1%)	0.361 ^b
Lymph node harvest [median (range)]	16 (3–39)	18 (6–43)	0.146 ^a	15 (7–32)	15 (0–35)	0.982 ^a
Length of hospital stay [median (range)] in days	6 (3–50)	6 (3–48)	0.838 ^a	8 (2–37)	6 (3–43)	0.436 ^a
Conversion	0 (0%)	0 (0%)		1 (2.1%)	0 (0%)	1.000 ^b
Reoperation	3 (7.0%)	2 (4.7%)	1.000 ^b	2 (4.3%)	1 (2.1%)	1.000 ^b
Readmission	7 (16.3%)	4 (9.3%)	0.520 ^b	10 (21.3%)	3 (6.4%)	0.035 ^b
Major complications (grade 3a–4b)	3 (7.0%)	2 (4.7%)	0.408 ^c	6 (12.8%)	3 (6.4%)	0.246 ^c
Mortality (grade 5)	0 (0%)	0 (0%)	1.000 ^b	1 (2.1%)	0 (0%)	1.000 ^b

^aMann–Whitney *U* test^bFisher's exact test^cPearson χ^2 test**Table 3** Individual on-target cut-off values and inherent risk for continuous and binary quality indicators, respectively, for surgeon A and B used for cumulative sum charting, based on laparoscopic cases

	Surgeon A	Surgeon B
Operation time in minutes		
Mean \pm standard deviation	222 \pm 43	278 \pm 54
Allowable slack	10.6	13.4
On-target cut-off ^a	233	292
Number of harvested lymph nodes		
Mean \pm standard deviation	16.4 \pm 8.5	16.6 \pm 6.8
Allowable slack	2.1	1.7
On-target cut-off ^a	14	15
Length of stay in days		
Mean \pm standard deviation	9.3 \pm 9.0	9.3 \pm 6.3
Allowable slack	2.2	1.6
On target cut-off ^a	12	11
Clavien–Dindo 3b–5		
Inherent risk	7.0%	14.9%

^aRounded to whole numbers as clinically applied

experience, there is no apparent learning curve, as observed for surgeon A. Similar observations have been previously reported for other procedures where skills were transferable to a new technique [21]. CUSUM plots are a widespread tool for sequential quality control and learning curve assessment [22] and have already been used to assess various colorectal procedures [23]. To our knowledge, there is only one recent study by Yamaguchi et al. using CUSUM charts to assess the learning curve in robotic rectal surgery [9]. A minimum of 25 robotic procedures was reported

as necessary to achieve proficiency. However, this study included a variety of procedures with and without lateral lymph node dissection and for a range of indications. No special measures were taken to account for selection bias (due to case mix) and only a single quality indicator, namely 'operation time', was used in the CUSUM analysis. Besides this, the level of proficiency in laparoscopic rectal surgery prior to starting with robotics was not stated. In contrast, our study included a defined standard procedure for a narrow field of indications, and the proficiency targets were calculated based on non-learning curve laparoscopic TMEs which were matched to consecutive robotic cases to minimise selection bias. Although operation time is an important indicator for the learning curve, opinions on the appropriate choice of quality indicators differ and other indicators are likely equally important to assess safety and oncological adequacy. For many surgeons, important quality criteria are feasibility, safety and ergonomics. For the hospital management, important quality indicators might include operating time, associated costs and marketing effects. For the patient, postoperative pain, functional outcome, cosmesis and disease-free as well as overall survival are of likely most interest. For all of these quality indicators, CUSUM charts can be constructed. Operation time is often considered an important surrogate for the adoptability and competitiveness of a new technique. Consequently, to compensate for longer operation times, the other benefits of a new technique may need to be substantial. Although a longer operation time is still a common argument against robotic TME [24, 25], our experience does not support this even when including docking time. With increased experience of the theatre team, the set-up and docking time could

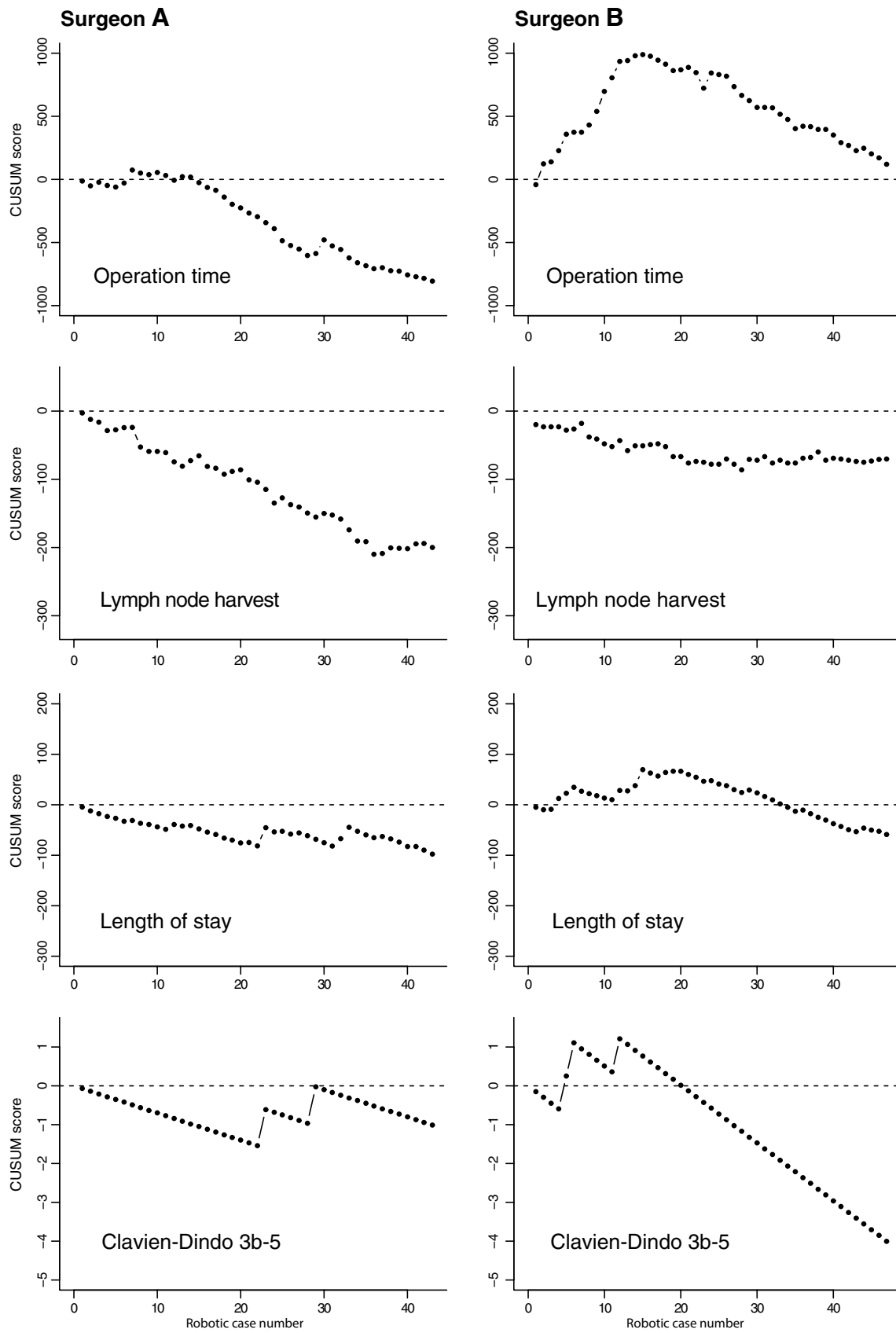


Fig. 2 Cumulative sum (CUSUM) charts for four quality indicators for the first consecutive robotic total mesorectal excisions performed by surgeon A and B

be reduced and allowed similar operation times as in laparoscopy after only 15 cases. It is important to mention that our study not only reflects the learning curve of the surgeon but also that of the whole team for whom the robotic platform was a novelty as well. The assistant standing at the patient's side plays an important role when it comes to active assistance via the auxiliary port, although the four arm robotic system with operator-guided camera as used in this study makes the console surgeon less dependent on the assistant's experience.

There was no statistically significant difference in the R0 resection rate between the laparoscopic and the robotic group. Although previous evidence has suggested similar circumferential margin positivity in open, laparoscopic and robotic TME, this was less clear during the learning curve phase [26]. In fact, surgeon B had less positive margins in robotic than in laparoscopic TME which is not entirely explained by differences of stage, site of tumour and neo-adjuvant treatment. This result is clinically relevant as it demonstrates non-inferiority of robotic TME even during the learning curve.

Lymph node harvest, often used as a surrogate marker for quality of surgery, was not adversely influenced during the learning period as shown in the corresponding CUSUM chart. However, lymph node count as a global quality indicator has been called into question as it does not only reflect the surgeon's effort alone but also pathological retrieval technique and tumour biology [27, 28].

The length of stay in hospital was prolonged in the first 15 cases during the learning curve of surgeon B, with a corresponding upward inflection in the CUSUM chart representing procedures not on target in this period. Length of stay may reflect the invasiveness of a lengthy procedure which may influence the degree of inflammatory response leading to a prolonged time to recover.

No perioperative death occurred in the robotic group and major complications were less common than in the laparoscopic group which was more apparent for the less-experienced surgeon B. For a new technique to be implemented successfully, it is crucial that quality indicators concerning patient safety are in-target as fast as possible. Due to the relatively static robotic platform, the lack of haptic feedback and the tunnel view, there is a potential for specific and severe complications in robotic surgery [29]. In particular, the tunnel view may be challenging for inexperienced colorectal surgeons who need more integral views to establish pattern recognition as provided in conventional laparoscopy. In our study, there was no conversion to open surgery in the robotic group. Similarly, a systematic review of short-term outcomes of robotic rectal cancer surgery found a lower conversion rate in the robotic than in the laparoscopic group [30]. Vascular dissection and medial to lateral mobilisation as well as the take down of the splenic

flexure can be challenging in both laparoscopic and robotic surgery. For the medial to lateral approach in robotic TME, it is crucial that the inferior mesenteric axis is identifiable before docking the robot. Experience in conventional laparoscopy helps efficiently use patient positioning in order to optimise exposure. The novice robotic surgeon may find this especially challenging as re-positioning of the patient is not possible once the robot is docked unless a synchronised operating table is used which was not the case in this study. In our institution, a standardised approach for complete mobilisation of the splenic flexure for both laparoscopic and robotic surgery has been developed which enables the surgeon to perform this complex step in a methodological manner resulting in a low conversion rate [11].

Addressing selection bias is important when evaluating a new technique as often straightforward cases (females, low BMI, small tumour size, no previous abdominal surgery, no preoperative radiotherapy) are chosen initially with the aim to progress to more complex cases with growing expertise. Matching laparoscopic cases to the consecutive robotic cases ensures that the corresponding laparoscopic reference cases theoretically could have been operated by robotics at the same stage of training. So, not only the robotic but also the matched laparoscopic cases were of increasing complexity and were well suited to define inherent risk and allowable slack.

Using our approach, CUSUM curves of different surgeons should not be compared without caution as the robotic performance is only benchmarked against an already achieved personal laparoscopic level of proficiency. If a surgeon had an inconsistent laparoscopic performance, this would result in a high allowable slack or inherent risk of quality indicators. Although being on target with reference to their own performance, they might be below standard when compared to other surgeons. On the other hand, a skilled and experienced surgeon who achieves consistent results leading to low standard deviations or inherent risks of quality indicators may find it difficult to reach a similar level when adopting a new technique. Our approach specifically answers the question as to whether it is worth changing from the current method, even when it produces good results. To prevent unnecessary harm to the patient, constant monitoring of the learning process is mandatory.

CUSUM curves for continuous quality indicators may be subject to systematic bias. The transition point of the CUSUM curve from off-target to on-target depends upon the allowable slack which may be somewhat arbitrary in its definition. The higher the allowable slack, the shorter the learning process becomes and vice versa. Half of a standard deviation is a commonly used slack when controlling industrial processes where the reference group is usually large leading to a relatively small standard deviation and a tight slack. In our model where only the surgeon's

own expert cases constituted the reference group, standard deviations were relatively high resulting in large allowable slacks when using half a standard deviation. Therefore, the tolerable slack was reduced to a quarter of the standard deviation from the mean in our approach.

Many expert laparoscopic colorectal surgeons are reluctant to start with robotic TMEs as they fear that they may have to repeat the learning process with all its negative implications. This study, however, provides evidence that the laparoscopic skill and procedural knowledge may be transferable to the robotic approach. The CUSUM charts also suggest that there may even be a “stepping down” from a more difficult laparoscopic to a less difficult robotic technique. This means that surgeons who are not able to perform a TME by laparoscopy might be able to do so by the robotic approach.

This study has limitations due to its retrospective nature, low case numbers, small number of surgeons included and the lack of quality grading of the specimens by the pathologist. The high number of cases proctored may also have a considerable impact on the learning curve, which could not be assessed. We had ten cases proctored for each surgeon, which may vary from other centres due to the cost implications of proctoring. Furthermore, the model used for this study only allows the assessment of a surgeon’s robotic performance in comparison to his own laparoscopic proficiency. Nevertheless, this study suggests that skills attained during laparoscopic TME surgery are transferable to robotic surgery and that the formal learning curve of experienced laparoscopic colorectal surgeons for robotic TME may be limited when compared to their individualised laparoscopic proficiency targets.

Acknowledgements The authors thank Sarah Marley, M.Sc., for her support in reviewing and providing statistical advice on this manuscript. Special thanks to Karen Flashman for data collection.

Compliance with ethical standards

Disclosures Manfred Odermatt, Jamil Ahmed, Sofoklis Panteleimonitis, Jim Khan and Amjad Parvaiz have no conflicts of interest or financial ties to disclose.

References

- Shah J, Vyas A, Vyas D (2014) The History of Robotics in Surgical Specialties. *Am J Robot Surg* 1:12–20
- Feroci F, Vannucchi A, Bianchi PP, Cantafio S, Garzi A, Formisano G, Scatizzi M (2016) Total mesorectal excision for mid and low rectal cancer: Laparoscopic vs robotic surgery. *World J Gastroenterol* 22:3602–3610
- D’Annibale A, Pernazza G, Monsellato I, Pende V, Lucandri G, Mazzocchi P, Alfano G (2013) Total mesorectal excision: a comparison of oncological and functional outcomes between robotic and laparoscopic surgery for rectal cancer. *Surg Endosc* 27:1887–1895
- Bianchi PP, Ceriani C, Locatelli A, Spinoglio G, Zampino MG, Sonzogni A, Crosta C, Andreoni B (2010) Robotic versus laparoscopic total mesorectal excision for rectal cancer: a comparative analysis of oncological safety and short-term outcomes. *Surg Endosc* 24:2888–2894
- Araujo SE, Seid VE, Klajner S (2014) Robotic surgery for rectal cancer: current immediate clinical and oncological outcomes. *World J Gastroenterol* 20:14359–14370
- Wang M, Kang X, Wang H, Guan W (2014) A meta-analysis on the outcomes and potential benefits of hybrid robotic technique in rectal cancer surgery. *Zhonghua Wei Chang Wai Ke Za Zhi* 17:785–790
- Trastulli S, Farinella E, Cirocchi R, Cavaliere D, Avenia N, Sciannameo F, Gulla N, Noya G, Boselli C (2012) Robotic resection compared with laparoscopic rectal resection for cancer: systematic review and meta-analysis of short-term outcome. *Colorectal Dis* 14:e134–e156
- Collinson FJ, Jayne DG, Pigazzi A, Tsang C, Barrie JM, Edlin R, Garbett C, Guillou P, Holloway I, Howard H, Marshall H, McCabe C, Pavitt S, Quirke P, Rivers CS, Brown JM (2012) An international, multicentre, prospective, randomised, controlled, unblinded, parallel-group trial of robotic-assisted versus standard laparoscopic surgery for the curative treatment of rectal cancer. *Int J Colorectal Dis* 27:233–241
- Yamaguchi T, Kinugasa Y, Shiomi A, Sato S, Yamakawa Y, Kagawa H, Tomioka H, Mori K (2015) Learning curve for robotic-assisted surgery for rectal cancer: use of the cumulative sum method. *Surg Endosc* 29:1679–1685
- 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
- Ahmed J, Kuzu MA, Figueiredo N, Khan J, Parvaiz A (2016) Three-step standardized approach for complete mobilization of the splenic flexure during robotic rectal cancer surgery. *Colorectal Dis* 18:O171–O174
- Bolsin S, Colson M (2000) The use of the Cusum technique in the assessment of trainee competence in new procedures. *Int J Qual Health Care* 12:433–438
- Chang WR, McLean IP (2006) CUSUM: a tool for early feedback about performance? *BMC Med Res Methodol* 6:8
- Kestin IG (1995) A statistical approach to measuring the competence of anaesthetic trainees at practical procedures. *Br J Anaesth* 75:805–809
- Montgomery DC (2009) Introduction to statistical quality control, 6th edn. Wiley, Hoboken
- Thoemmes F (2012) Propensity score matching in SPSS. <https://arxiv.org/ftp/arxiv/papers/1201/1201.6385.pdf>. arXiv:1201.6385; Accessed 18 Aug 2016.
- R Core Team (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Ho DE, Imai K, King G, Stuart EA (2011) MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw* 42:28
- Bowers J, Fredrickson M, Hansen B (2010) RIttools: Randomization Inference Tools. pp R package
- Iacus S, King G, Porro G (2009) cem: Software for Coarsened Exact Matching. pp R package
- Colquhoun PH (2008) CUSUM analysis of J-pouch surgery reflects no learning curve after board certification. *Can J Surg* 51:296–299
- Biau DJ, Resche-Rigon M, Godiris-Petit G, Nizard RS, Porcher R (2007) Quality control of surgical and interventional procedures: a review of the CUSUM. *Qual Saf Health Care* 16:203–207

23. Bowles TA, Watters DA (2007) Time to CUSUM: simplified reporting of outcomes in colorectal surgery. *ANZ J Surg* 77:587–591
24. Trinh BB, Jackson NR, Hauch AT, Hu T, Kandil E (2014) Robotic versus laparoscopic colorectal surgery. *JSLS* 18:3
25. Fung AK, Aly EH (2013) Robotic colonic surgery: is it advisable to commence a new learning curve? *Dis Colon Rectum* 56:786–796
26. de Jesus JP, Valadao M, de Castro Araujo RO, Cesar D, Linhares E, Iglesias AC (2016) The circumferential resection margins status: A comparison of robotic, laparoscopic and open total mesorectal excision for mid and low rectal cancer. *Eur J Surg Oncol* 42:808–812
27. Baxter NN (2009) Is lymph node count an ideal quality indicator for cancer care? *J Surg Oncol* 99:265–268
28. Wang J, Kulaylat M, Rockette H, Hassett J, Rajput A, Dunn KB, Dayton M (2009) Should total number of lymph nodes be used as a quality of care measure for stage III colon cancer? *Ann Surg* 249:559–563
29. Cooper MA, Ibrahim A, Lyu H, Makary MA (2015) Under-reporting of robotic surgery complications. *J Healthc Qual* 37:133–138
30. Scarpinata R, Aly EH (2013) Does robotic rectal cancer surgery offer improved early postoperative outcomes? *Dis Colon Rectum* 56:253–262