Original article

doi:10.1111/codi.15269

European expert consensus on a structured approach to training robotic-assisted low anterior resection using performance metrics

S. Tou*, M. Gómez Ruiz†, A. G. Gallagher‡, K. E. Matzel§ and the ESCP ASPIRE collaborative[1]

*Department of Colorectal Surgery, Royal Derby Hospital, University Hospitals of Derby and Burton NHS Foundation Trust, Derby, UK, †Cirugía Colorrectal – Cirugía General y del Aparato Digestivo, Hospital Universitario Marqués de Valdecilla, Santander, Spain, ‡ORSI Academy, Melle, Belgium, and §Section of Coloproctology, Department of Surgery, University of Erlangen-Nürnberg, FAU, Erlangen, Germany

Received 18 April 2020; revised 30 June 2020; accepted 2 July 2020; Accepted Article online 14 July 2020

Abstract

Aim The aim was to develop and operationally define 'performance metrics' that characterize a reference approach to robotic-assisted low anterior resection (RA-LAR) and to obtain face and content validity through a consensus meeting.

Method Three senior colorectal surgeons with robotic experience and a senior behavioural scientist formed the Metrics Group. We used published guidelines, training materials, manufacturers' instructions and unedited videos of RA-LAR to deconstruct the operation into defined, measurable components – performance metrics (i.e. procedure phases, steps, errors and critical errors). The performance metrics were then subjected to detailed critique by 18 expert colorectal surgeons in a modified Delphi process.

Results Performance metrics for RA-LAR had 15 procedure phases, 128 steps, 89 errors and 117 critical errors in women, 88 errors and 118 critical errors in men. After the modified Delphi process the final performance metrics consisted of 14 procedure phases, 129 steps, 88 errors and 115 critical errors in women, 87 errors and 116 critical errors in men. After discussion by the Delphi panel, all procedure phases received unanimous consensus apart from phase I (patient positioning and preparation, 83%) and phase IV (docking, 94%).

Conclusion A robotic rectal operation can be broken down into procedure phases, steps, with errors and critical errors, known as performance metrics. The face and content of these metrics have been validated by a large group of expert robotic colorectal surgeons from Europe. We consider the metrics essential for the development of a structured training curriculum and standardized procedural assessment for RA-LAR.

Keywords metrics, training, proficiency-based progression, low anterior resection, robotic surgery, colorectal surgery

What does this paper add to the literature?
The present study is the first to describe a structured approach to develop performance metrics for training in robotic-assisted low anterior resection for rectal cancer.

Introduction

For the past decades, we have been in search of better outcomes for patients undergoing colorectal cancer surgery. From optimizing surgical techniques, the use of adjuvant and neoadjuvant therapy, the adoption of a minimally invasive approach to enhanced recovery after surgery [1], all these elements are incorporated into the principle of aggregation of marginal gains, which was popularized by Sir Dave Brailsford [2]. Furthermore, for more complex and lower volume procedures, there is evidence to encourage centralization to enhance surgeons' and their team's experience as it may help to improve surgical outcomes and survival [3].

Centralization of rectal cancer surgery improves long-term survival [4]. However, this volume and

Correspondence to: S. Tou MS, FRCS, M.MinInvSu, Consultant Colorectal Surgeon and Honorary Associate Professor, University Hospitals of Derby and Burton NHS Foundation Trust, Derby DE22 3NE, UK.
E-mail: samsontou@aol.com

[1]ESCP ASPIRE collaborative members are listed in the Acknowledgements.

outcome relationship may not always be directly correlated, and volume alone cannot be an indicator to assure good patient outcomes [5,6]. Recent research into surgical skills has identified a strong correlation between surgeons' technical performance, patients' complication rates and clinical outcomes [7,8]. Therefore, the focus has now turned to implementing optimal surgical training and improving surgeons' skills.

The European School of Coloproctology (ESC) of the European Society of Coloproctology (ESCP) was set up to benchmark surgical training at a European level, and the robotic colorectal surgery curriculum is one of the training programmes being established [9]. In the search for a more objective and transparent approach to surgical training, we have identified a scientific approach, known as proficiency-based progression (PBP) training [10–12]. The core of this approach is to deconstruct a surgical procedure to operational defined performance metrics [10,11,13–17]. These metrics can facilitate training and assessment and allow learners to progress in their training based on their proficiency, rather than the number of cases performed or duration of practice. This method has been applied to specialities including orthopaedics [15,17], cardiology [14,18], anaesthetics [16,19] and interventional neuroradiology [13,20]. PBP training has been shown to significantly reduce intra-operative errors (> 40%) [18,19,21–25]. Therefore, the ESC aims to introduce this scientific training approach in the robotic colorectal surgery training programme.

Low anterior resection (LAR) for rectal cancer is an index operation in the field of coloproctology which is commonly performed with robotic-assisted (RA) surgery. An optimally performed RA-LAR potentially minimizes adverse clinical, oncological and functional outcomes for patients undergoing this procedure. Our study aimed to develop and objectively define performance 'metrics' that characterize a reference approach to RA-LAR for rectal cancer, and to obtain face and content validity through a consensus meeting (i.e. with a Delphi panel) of very experienced and expert robotic colorectal surgeons.

Method

The principle of metric development and stress testing (face and content validation) for PBP training has been described in non-colorectal specialities previously [10,11,13–17]. This principle was applied when developing the RA-LAR metrics, and is described below.

Metrics team

Approval of the conduct of the study was granted by the Ethics Committee at the Instituto de Investigacion Marques de Valdecilla (IDIVAL), Santander, Cantabria, Spain (Código interno: 2020.026). The Metrics Group consists of three experienced colorectal surgeons (KM, MGR, ST) with special interest in robotic surgery, who are involved in setting up the ESCP Colorectal Robotic Surgery (ESCP CRS) Programme, a senior behavioural scientist and an education-training expert (AGG).

RA-LAR metrics development

A detailed task analysis and deconstruction process were used to break down a 'reference' approach to the performance of the complete RA-LAR for rectal cancer procedure in small non-overlapping parts [10,11,14,17]. Published written guidelines, video teaching materials, manufacturers' instructions and access of more than 10 anonymized unedited RA-LAR videos performed by surgeons with different levels of experience supported the metrics development process. The goal was to characterize a reference approach to RA-LAR procedures for rectal cancer in female and male patients. A reference procedure is assumed to be straightforward and uncomplicated to guide trainees to learn the 'recipe' to perform these procedures, before performing more complex operations such as patients with high body mass index or with locally advanced disease. The phases and steps are the same between women and men undergoing RA-LAR apart from during the rectal dissection phase due to anatomical considerations.

A 1-day preliminary face-to-face planning meeting, three 1½-day face-to-face meetings for metrics identification and definition, and four 1½-day face-to-face meetings for the metric stress test were conducted. Video-conferencing using Zoom (San Jose, California, USA) and email exchange were used to complement face-to-face meetings for further clarification of the metrics development.

At the beginning of the metrics development, the Metrics Group agreed on some definitions as follows.

Performance metrics: Units of observable behaviour which together constitute a stepwise description of a reference approach to a procedure.

Procedure phase: A group or series of integrally related events or actions that, when combined with other phases, make up or constitute a complete operative procedure.

Step: A component task, the series aggregate of which forms the completion of a specific procedure.

Error: A deviation from optimal performance.

Critical error: Major deviation from optimal performance which has a likelihood of causing harm to the patient or compromising the safe completion of the procedure [13,15,17].

The metrics therefore consist of procedural phases involved in an RA-LAR. Each phase comprises specific steps required to accomplish that phase. The importance of the metrics approach in defining these phases and steps is that they are clear and unambiguous. The step either occurred or did not occur, and can be scored as such by an external reviewer with a high degree of reliability [26–28]. Similarly, errors and critical errors were defined associated with particular steps within different phases of the procedure. For errors, behaviours exhibited by the operator may not necessarily in and of themselves lead to a bad outcome or an event with more serious consequences but their enactment sets the stage or increases the probability for a more serious event to occur or detracts from the efficient execution of the desired procedure. In contrast, a critical error is a more serious occurrence and represents operative performance that could either jeopardize the outcome of the procedure or lead to significant iatrogenic damage [13,14,17]. Figure 1 illustrates an example of a procedural phase characterized for the RA-LAR. The sequence of the procedure assessed using metrics in the RA-LAR is included in Appendix S1. In addition to the metrics, valuable knowledge and principles of the operation were compiled (such as applied anatomy) to facilitate the learning process; these formed the didactic component for the learner during the training process.

Once the Metrics Group defined the metrics, they were then used to score four unedited anonymized RA-LARs performed by four different surgeons. The scorings were performed by the members of the Metrics Group independently. Any difference in the scoring was discussed in order to identify discrepancies in interpretation or ambiguities in the metric definition. Based on this process and if agreed, changes were made in the metrics to facilitate the scoring agreement. This process was repeated for each video until the Metrics Group was satisfied with the metrics and they could be scored with a high degree of reliability (i.e. inter-rater reliability > 0.8 – the internationally agreed gold standard) [29,30]. In the end, there were 15 phases, 128 steps, 89 errors and 117 critical errors in women, 88 errors and 118 critical errors in men.

Metrics stress testing (face and content validation) with a modified Delphi approach

Once the metrics for the RA-LAR were defined and characterized, face validity and content were verified by a group of experienced robotic colorectal surgeons. To provide a more objective and independent assessment of the metrics an international panel of expert minimally invasive colorectal surgeons with robotic experience were invited to join the Delphi panel [13–17].

Eighteen expert colorectal surgeons including the three Metrics Group members from 10 countries, a non-voting behavioural scientist and a senior administrative member for the ESCP CRS Working Group attended a consensus meeting in Munich, 8 February 2020 (Table 1). The panel was chosen for their colorectal robotic surgical experience and their demonstrated educational interests and commitment. The age of the panel ranged from 40 to 62 years; there were three female surgeons. Ten panel members were head of their respective departments, and five were full professors affiliated with universities. The combined robotic colorectal procedures performed or supervised by the Delphi panel were more than 4000.

A brief overview of the project and meeting objectives was presented. Background information regarding proficiency-based training, prior literature demonstrating the validity of this training approach for procedural specialties, and the specific objectives of the current Delphi panel were reviewed [31]. Each phase of the procedure, the procedural steps that were included in that phase, and the potential errors were presented. It was also explained that the associated metrics had been developed by the Metrics Group for a reference approach to RA-LAR for rectal cancer. It was acknowledged that the designated reference procedure might not reflect the exact techniques employed by individual Delphi panellists, but that the operative steps presented accurately embodied the essential and key components of the procedure and 'were not wrong' [13–15,17].

To assess the correlation of the procedure steps, errors and critical errors before and after the Delphi process, changes were analysed with the Pearson chi-squared test (IPM SPSS Statistics for Windows, version 26, IBM Corp., Armonk, New York, USA). A P value of < 0.05 was considered statistically significant.

Results

The Metrics Group proposed 15 phases included in the RA-LAR, and each phase had a defined beginning and end. During the Delphi process, two phases (stoma formation and wound closure) were combined, to allow accurate representation of the sequence of the steps during surgery. Therefore, there was a total of 14 phases in the final RA-LAR procedure (Table 2).

In the end, there are 14 phases, 129 steps performed during a RA-LAR; one step was added (sterile draping of the cart in phase II), and 16 steps were modified during the Delphi meeting (Table 3).

(a)

VII. Splenic flexure mobilization (*B, atraumatic instrument to retract the descending mesocolon; E, left mesotransverse colon is completely mobilised)*	Step	Error	Critical Error	DNTT
54. Traction on the descending mesocolon antero-laterally and exposing the clip of the mesenteric vein	x			
55. Continue medial to lateral dissection, to inferior boarder of the pancreas, and laterally continues until the visualisation of the posterior part of descending/sigmoid colon	x			
56. Assistant's forceps lift the transverse measocolon opening the space between transverse mesocolon and the body and taiol of the pancreas				
57. Entering the lesser sac, ventral to the body of the pancreas and lateral to the stump of IMV	x			
58. Detach the transverse mesocolon medial to laterally towards the tail of the pancreas	x			
59. Traction on the descending mesocolon medially and lateral counter-traction of the peritoneum	x			
60. Dissection along the lateral attachment of the descending colon, working cranially towards the splenic flexure	x			
61. Detach greater omentum from the colon at the level of splenic flexure	x			
62. Using traction and counter-traction to expose the attachment of the omentum to the transverse colon, and detach and separate (either lateral to medial, or medial to lateral would be acceptable)	x			
63. Left transverse mesocolon is mobilised but not beyond the middle colic vessels (preserving the vessels)	x			

(b)

	Step	Error	Critical Error	DNTT
• *Failure to maintain position of port E*				
• *Dissection was not in the embryological plane E*		x		
• *Wrong selection of instruments, e.g. traumatic instrument on tissue E*				
• *Wrong aaplication of instruments e.g. on bowel wall and vessels E*				
• *Use of energy device not under direct vision CE*			x	
• *Injury to small or large bowel CE*				
• *Injury to vessels CE*				
• *Injury to measocolon CE*				
• *Instrument positioning, inadequate exposure E*				
• *Injury to pancreas CE*				
• *Injury to spleen CE*				
• *Injury to mesenteric arcade CE*				
• *Injury to stomach CE*				
• *Injury to middle colic vessels CE*				
• *Injury to splenic artery CE*				
• *Injury to splenic vein CE*				

Figure 1 Example of a phase during robotic-assisted low anterior resection that was characterized with steps, errors and critical errors. DNTT, damage to non-target tissue.

Table 1 Number of surgeons from each country represented in the Delphi panel.

Country	Number of surgeons
France	1
Spain	2
UK	5
Germany	3
Belgium	1
Italy	2
Poland	1
Ireland	1
Sweden	1
The Netherlands	1
Total	18

There were 89 procedure errors identified for RA-LAR in female patients and 88 in male patients. After the Delphi meeting, the final number of procedure errors in women undergoing RA-LAR was 88, and 87 in men (Table 4).

Similarly, small changes were made in the Delphi meeting, and the numbers of procedure critical errors identified for RA-LAR were 115 and 116, respectively, for women and men (Table 5). Furthermore, the number of procedure steps, errors and critical errors before and after the Delphi changes were highly correlated (Pearson correlation coefficient for procedure steps $r = 0.999$, $P < 0.001$; errors $r = 0.98$, $P < 0.001$; and critical errors $r = 0.985$, $P < 0.001$).

Table 6 shows the consensus from the Delphi panel in each procedure phase after incorporated changes during the meeting. Apart from the patient positioning and preparation, and docking of the robot, other phases received unanimous votes from the panel.

Discussion

In this study, the performance metrics (procedure phases, steps, errors, critical errors) for RA-LAR for rectal cancer could be characterized. This was achieved through a systematic and structured approach. A Metrics Group was formed with three expert colorectal surgeons with a particular interest in robotic surgery, involved in setting up a training curriculum at a European level (ESC), and a senior behavioural scientist who has more than two decades of experience in surgical training. The development process of the metrics was comprehensive and involved reviewing the published guidelines, training materials, manufacturers' instructions and unedited anonymized videos of RA-LAR performed by surgeons with different levels of experience. The metrics that have been developed are not confined to any specific robotic platform.

The metrics were also subjected to stress testing. First, the individual surgeons in the Metrics Group independently scored unedited videos of RA-LAR performed by different surgeons. Differences in scoring resulted in further discussion and refinement of the metrics operational definitions to facilitate consensus and reliable scoring. Second, the metrics produced by the Metrics Group were then scrutinized by a panel of expert colorectal surgeons (18, from 10 different countries). The expert panel during the Delphi meeting provided an opportunity to optimize the metrics further. The Delphi panel understood that surgeons might work differently, but the goal was to achieve a standardized structure approach to a 'reference' RA-LAR. The metrics presented in the Delphi meeting may not be in the same way or sequence that each surgeon is practising, but the metrics are not wrong and should be suitable for learners. The pre- and post-Delphi metrics are highly correlated (Tables 2–5). After incorporating the changes suggested by the Delphi panel voting was obtained at the end of the discussion of each phase. All the phases received unanimous agreement apart from phase I, Patient position and preparation, and phase IV, Docking. The main discussion in phase I was about the padding of the patient, their face, body and limb protection to ensure the safety of the patient. There was no suggestion to change the metrics, but the knowledge of compartment syndrome should be added to the didactic part of the training programme. There was no issue raised with regard to the docking phase of the RA-LAR.

With the evolution of surgical training, we have witnessed a shifting paradigm of the training curriculum from the apprenticeship model to a more organized approach such as the UK LapCo programme [32]. However, it is not straightforward to determine when the learners can progress to the next phase of training and whether this is determined based on volume or duration. The volume–outcome relationship has attracted much attention in the procedural-based intervention [5,33]. While it is expedient to measure volume, the quality of the procedure and volume relationship may not always be straightforward [6]. One of the primary functions of the types of performance metrics reported on in this study is to accelerate the learning curve. They provide performance feedback to the trainee which affords deliberate practice rather than repeated practice [34]. This means that the trainee receives feedback which is proximate to their performance which informs them what they did wrong and how it should optimally be performed [35].

Table 2 The beginning and end of the different phases of the reference approach to robotic-assisted low anterior resection and the changes agreed and voted on by the Delphi panel.

Procedure phase	Title	Phase – begins	Phase – ends
I	Patient positioning and preparation	Completion of WHO checklist	Patient is on the table before prepping
II	Preparation of operative field	Creation of a sterile field	Patient is draped
III	Trocar position	Incision/insertion of trocars	Removal of laparoscopic instruments
IV	Docking	Advance the patient cart to the patient	Operating surgeon takes control at the console
V	IMA dissection/ligation	Visualize the working end of all three robotic instruments intraabdominally	Complete division of the IMA
VI	IMV exposure and ligation	Atraumatic instrument to retract the descending mesocolon	Complete division of the IMV
VII	Splenic flexure mobilization	Atraumatic instrument to retract the descending mesocolon	Left mesotransverse colon is completely mobilized
VIII	Complete mobilization of left colon	Atraumatic instrument to retract the descending mesocolon	Release the lateral attachment of the sigmoid colon
IX	Rectal dissection/rectal transection (TME/LAR) – separately for female/male patient	Visualize the working end of all three robotic instruments in the pelvis	The divided rectum is placed in the abdominal cavity and in view
X	Undocking the system	Robotic instruments removed	Patient cart removed
XI	*Specimen extraction and re-establishing the pneumoperitoneum (adjust the position of the table)*	Make Pfannenstiel incision	Permanently or temporarily closing the transverse incision to re-establish the pneumoperitoneum
XII	Anastomosis	Move the proximal bowel to pelvis	Check the anastomosis for leakage, e.g. air leak test, rigid or flexible sigmoidoscopy
XIII	*Stoma formation and wound closure*	*Incision at the stoma site*	*Application of stoma bag/appliance*
XIV	Transfer patient from operating table to bed	Transfer patient to bed	Patient out of the operating room
$N = 14$	*Stoma formation and wound closure were combined to one phase during Delphi meeting*		

Changes are in italic.
IMA, inferior mesenteric artery; IMV, inferior mesenteric vein; LAR, low anterior resection; TME, total mesorectal excision; WHO, World Health Organization.

Furthermore, these types of metrics are the foundation of effective virtual reality simulation training [11,13,18,25,36]. Explicitly defined procedure performance metrics, scored in a binary fashion, are more difficult to develop than off-the-shelf Likert-type scales (e.g. OSATS or GEARS) [37,38]. They are more reliable, however (e.g. OSATS and GEARS) [26,28] and provide the trainee with objective, transparent and explicit feedback which is derived from a consensus of expert senior surgeons [11–13,15,17]. In contrast to hand/tool 'efficiency' measures these types of performance metrics will also underpin and bolster the development of effective virtual reality simulation and machine learning [20,39–41].

The approach used to characterize an RA-LAR in this study provides structured and objective metrics that are explicit and transparent and allow learners to focus on achieving the steps required before progress to the next level – PBP training. This PBP training has been utilized in different specialities and has been shown to significantly reduce intra-operative errors [18,19,21–25,42]. Furthermore, the metrics will serve as a template for trainers to provide structured training for a surgical procedure. The approach also provides effective

Table 3 Steps before and after the Delphi meeting.

Procedure phase	Title	Steps before Delphi	Steps after Delphi	Added	Deleted	Modified
I	Patient positioning and preparation	11	11	0	0	1
II	Preparation of operative field	5	6	1	0	1
III	Trocar position	16	16	0	0	3
IV	Docking	7	7	0	0	2
V	IMA dissection/ligation	7	7	0	0	1
VI	IMV exposure and ligation	6	6	0	0	0
VII	Splenic flexure mobilization	10	10	0	0	0
VIII	Complete mobilization of left colon	4	4	0	0	0
IX	Rectal dissection/rectal transection (TME/LAR) – separately for female/male patient	19	19	0	0	0
X	Undocking the system	3	3	0	0	1
XI	Specimen extraction and re-establishing the pneumoperitoneum (adjust the position of the table)	7	7	0	0	0
XII	Anastomosis	15	15	0	0	1
XIII	Stoma formation and wound closure	16	16	0	0	6
XIV	Transfer patient from operating table to bed	2	2	0	0	0
N = 14		128	129	1	0	16

IMA, inferior mesenteric artery; IMV, inferior mesenteric vein; LAR, low anterior resection; TME, total mesorectal excision.

Table 4 Errors before and after the Delphi meeting.

Procedure phase	Title	Errors before Delphi	Errors after Delphi	Added	Deleted	Modified
I	Patient positioning and preparation	3	3	0	0	0
II	Preparation of operative field	2	2	0	0	0
III	Trocar position	8	8	0	0	0
IV	Docking	6	5	0	1	0
V	IMA dissection/ligation	9	11	2	0	0
VI	IMV exposure and ligation	9	10	1	0	0
VII	Splenic flexure mobilization	5	5	0	0	0
VIII	Complete mobilization of left colon	5	5	0	0	0
IX	Rectal dissection/rectal transection (TME/LAR) – separately for female/male patient	Women – 13 Men – 12	Women – 13 Men – 12	Women – 1 Men – 1	Women – 1 Men – 1	Women – 0 Men – 0
X	Undocking the system	3	3	0	0	0
XI	Specimen extraction and re-establishing the pneumoperitoneum (adjust the position of the table)	8	7	0	1	0
XII	Anastomosis	6	5	1	2	0
XIII	Stoma formation and wound closure	12	11	0	1	0
XIV	Transfer patient from operating table to bed	0	0	0	0	0
N = 14		Women – 89 Men – 88	Women – 88 Men – 87	Women – 5 Men – 5	Women – 6 Men – 6	Women – 0 Men – 0

IMA, inferior mesenteric artery; IMV, inferior mesenteric vein; LAR, low anterior resection; TME, total mesorectal excision.

Table 5 Critical errors before and after the Delphi meeting.

Procedure phase	Title	Critical errors before Delphi	Critical errors after Delphi	Added	Deleted	Modified
I	Patient positioning and preparation	13	13	0	0	2
II	Preparation of operative field	5	5	0	0	1
III	Trocar position	7	7	0	0	0
IV	Docking	3	4	1	0	0
V	IMA dissection/ligation	14	14	0	0	0
VI	IMV exposure and ligation	16	15	0	1	1
VII	Splenic flexure mobilization	11	11	0	0	0
VIII	Complete mobilization of left colon	7	7	0	0	0
IX	Rectal dissection/rectal transection (TME/LAR) – separately for female/male patient	Women – 12 Men – 13	Women – 13 Men – 14	Women – 1 Men – 1	Women – 0 Men – 0	Women – 0 Men – 0
X	Undocking the system	1	1	1	1	1
XI	Specimen extraction and re-establishing the pneumoperitoneum (adjust the position of the table)	8	7	0	1	1
XII	Anastomosis	11	9	0	2	0
XIII	Stoma formation and wound closure	7	7	0	0	0
XIV	Transfer patient from operating table to bed	2	2	0	0	0
N = 14		Women – 117 Men – 118	Women – 115 Men – 116	Women – 3 Men – 3	Women – 5 Men – 5	**Women – 6** Men – 6

IMA, inferior mesenteric artery; IMV, inferior mesenteric vein; LAR, low anterior resection; TME, total mesorectal excision.

Table 6 Results from the Delphi meeting and consensus reached at the end of each phase.

Procedure phase	Title	Percentage consensus (%)
I	Patient positioning and preparation	83
II	Preparation of operative field	100
III	Trocar position	100
IV	Docking	94
V	IMA dissection/ligation	100
VI	IMV exposure and ligation	100
VII	Splenic flexure mobilization	100
VIII	Complete mobilization of left colon	100
IX	Rectal dissection/rectal transection (TME/LAR) – separately for female/male patient	100 (female TME), 100 (male TME) 100 (transection of rectum)
X	Undocking the system	100
XI	*Specimen extraction and re-establishing the pneumoperitoneum (adjust the position of the table)*	100
XII	Anastomosis	100
XIII	*Stoma formation and wound closure*	100
XIV	Transfer patient from operating table to bed	100

IMA, inferior mesenteric artery; IMV, inferior mesenteric vein; LAR, low anterior resection; TME, total mesorectal excision.

learning because the metrics underpin a 'deliberate practice' approach to training rather than the traditional repeated practice approach [34]. Equally important, performance assessment using these metrics will provide feedback to the trainee that is transparent, objective and unambiguous, disadvantages often associated with

Likert-type scale assessment [26,27,30]. A recently published consensus statement has highlighted the importance of standardization of robotic total mesorectal excision. This study provides a good starting point for training in robotic rectal surgery [43]. The metrics described by us have objectively measurable steps with defined errors and critical errors and will further enhance training through a 'deliberate practice' method.

The debate of the approach for cancer of the rectum is ongoing [44], but RA-LAR has been increasingly performed [45]. Given the adverse outcomes associated with the new technique of rectal cancer surgery [46], although outcomes from large multicentre trials under way are awaited (COLOR III, RESET) [47,48], the importance of a structured and quality assured approach to training is needed to safeguard patients' outcomes. This is one of the fundamental remits of the ESCP CRS Working Group within the ESC.

The next phase, to continue this present work with the performance metrics for RA-LAR, is to obtain construct validity, i.e. whether the metrics distinguish between the objectively assessed performances of novices and very experienced robotic colorectal surgeons. The results from the construct validity are vital as the validated metrics can be used to inform training and assessment, i.e. setting a benchmark in surgical training [9], using PBP approach.

There are several limitations to this study. The proposed metrics are for a standard straightforward (i.e. reference approach) RA-LAR for female and male patients. The aim is to provide learners with a structured stepwise approach. The Delphi panel was aware occasionally that there are deviations from a 'reference' operation, and a slight variation of techniques may be needed. Equally, there is more than a single approach for a particular part of an operation. For example, the splenic flexure mobilization technique characterized in the metrics employs a supra-pancreatic approach, commonly practised in Europe. The learners will probably be exposed to other approaches once they have mastered the recommended technique described in these metrics.

The number of errors and critical errors is not exhaustive, but these are considered by the Metrics Group and the Delphi panel to be important in RA-LAR. There are no clinical outcomes to correlate with the scoring of these metrics. Therefore, it is not known at this stage which part of the metrics are more important than others, i.e. they are not weighted.

In conclusion, a commonly performed robotic rectal operation can be broken down to defined phases and steps and have measurable errors and critical errors, known as performance metrics. Evidence of the face and content of these metrics has been validated by a large group of expert robotic colorectal surgeons from Europe. Further development of these metrics is essential to guide the training curriculum and assessment for RA-LAR.

Acknowledgements

We thank Linda Ensinger, Vanessa McCourt and Amy Warner for providing support and preparation of the Delphi meeting.

European Society of Coloproctology – Advancing Standards and Practice in Robotic Surgical Education – the ESCP Aspire Collaborative: Shwan Amin (Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK); Paolo Bianchi (Ospedale della Misericordia, Grosseto, Italy); Rogier Crolla (Amphia Ziekenhuis, Breda, The Netherlands); Roland Croner (Universitat Magdeburg, Magdeburg, Germany); Nicola Eardley (Countess of Chester Hospital NHS Foundation Trust, Chester, UK); Eloy Espin-Basany* (Hospital Valle de Hebron, Barcelona, Spain); Charlie Evans (University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK); Carmen Gagigas Fernández (Hospital Universitario Marques de Valdecilla, Santander, Spain); Anthony Gallagher (ORSI Academy, Melle, Belgium); Roger Gerjy (Danderyd Hospital, Stockholm, Sweden); Marcos Gomez Ruiz (Hospital Universitario Marques de Valdecilla, Santander, Spain); Dieter Hahnloser* (Lausanne University Hospital, Lausanne, Switzerland); Jim Khan (Portsmouth Hospitals NHS Trust, Portsmouth, UK); Bernard Lelong (Comprehensive Cancer Centre, Marseille, France); Klaus Matzel (University Erlangen-Nurnberg, Erlangen, Germany); Paul Neary (Tallaght University Hospital, Dublin, Ireland); Daniel Perez (University Hospital Eppendorf, Hamburg, Germany); Giuseppe Spinoglio (Alessandria National Hospital, Alessandria, Italy); Samson Tou (University Hospitals of Derby and Burton NHS Foundation Trust, Derby, UK); Ellen van Eetvelde (University Hospital Brussels, Brussels, Belgium); Marek Zawadzki (Wroclaw Regional Hospital, Wroclaw, Poland).

*These are members of the collaborative but were unable to attend the Delphi meeting.

Conflicts of interest

ST received education grants from Intuitive Surgical and Medtronic. MGR received education grants from Intuitive Surgical and Medtronic and currently is Medical Advisor to Intuitive Surgical, Medtronic and J&J. AGG holds education research grants from Medtronic

(Dublin, Ireland), AO Education Institute (Davos, Switzerland) and the Arthroscopic Association of North America (Chicago, USA) for the investigation of metric-based education and training. KEM is Medical Advisor to Medtronic and Boehringer Ingelheim.

Funding

The study is part of the ongoing project for the Colorectal Robotic Surgery Working Group, European School of Coloproctology, European Society of Coloproctology. Medtronic provided the educational grant for this study but did not influence in selecting the experts, design and conduct of the research, data collection, analysis and preparation of the manuscript.

References

1 Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet* 2019; **394:** 1467–80.
2 Parker M. The aggregation of marginal gains. *Ann R Coll Surg Engl (Suppl)* 2011; **93:** 236–7.
3 Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? – The empirical relationship between surgical volume and mortality. *N Engl J Med* 1979; **301:** 1364–9.
4 Khani MH, Smedh K. Centralization of rectal cancer surgery improves long-term survival. *Colorectal Dis* 2010; **12:** 874–9.
5 Ibrahim AM, Ghaferi AA, Thumma JR, Dimick JB. Variation in outcomes at bariatric surgery centers of excellence. *JAMA Surg* 2017; **152:** 629–36.
6 Gallagher AG, Angelo RL, Kearney P. Factors associated with variation in outcomes in bariatric surgery centers of excellence. *JAMA* 2018; **320:** 1386–7.
7 Birkmeyer JD, Finks JF, O'Reilly A *et al.* Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013; **369:** 1434–42.
8 Curtis NJ, Foster JD, Miskovic D *et al.* Association of surgical skill assessment with clinical outcomes in cancer surgery. *JAMA Surg* 2020; **155:** 590–8.
9 Gomez Ruiz M, Tou S, Matzel KE. Setting a benchmark in surgical training – robotic training under the European School of Coloproctology, ESCP. *Colorectal Dis* 2019; **21:** 489–90.
10 Gallagher A. Metric-based simulation training to proficiency in medical education – what it is and how to do it. *Ulster Med J* 2012; **81:** 107–13.
11 Gallagher AG, O'Sullivan GC. *Fundamentals of Surgical Simulation; Principles and Practices.* London: Springer Verlag, 2011. https://www.springer.com/gp/book/9780857297624 (accessed May 2020).
12 Gallagher AG, Ritter EM, Champion H *et al.* Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005; **241:** 364–72.
13 Crossley R, Liebig T, Holtmannspoetter M *et al.* Validation studies of virtual reality simulation performance metrics for mechanical thrombectomy in ischemic stroke. *J Neurointerv Surg* 2019; **11:** 775–80.
14 Mascheroni J, Mont L, Stockburger M *et al.* International expert consensus on a scientific approach to training novice cardiac resynchronization therapy implanters using performance quality metrics. *Int J Cardiol* 2019; **289:** 63–9.
15 Kojima K, Graves M, Taha W, Cunningham M, Joeris A, Gallagher AG. AO international consensus panel for metrics on a closed reduction and fixation of a 31A2 pertrochanteric fracture. *Injury* 2018; **49:** 2227–33.
16 Ahmed OM, D O'Donnell B, Gallagher AG, Shorten GD. Development of performance and error metrics for ultrasound-guided axillary brachial plexus block. *Adv Med Educ Pract* 2017; **8:** 257–63.
17 Angelo RL, Ryu RK, Pedowitz RA, Gallagher AG. Metric development for an arthroscopic Bankart procedure: assessment of face and content validity. *Arthroscopy* 2015; **31:** 1430–40.
18 Cates CU, Lönn L, Gallagher AG. Prospective, randomised and blinded comparison of proficiency-based progression full-physics virtual reality simulator training versus invasive vascular experience for learning carotid artery angiography by very experienced operators. *BMJ Simul Technol Enhanc Learn* 2016; **2:** 1–5.
19 Srinivasan KK, Gallagher A, O'Brien N *et al.* Proficiency-based progression training: an 'end to end' model for decreasing error applied to achievement of effective epidural analgesia during labour: a randomised control study. *BMJ Open* 2018; **8:** e020099.
20 Liebig T, Holtmannspötter M, Crossley R *et al.* Metric-based virtual reality simulation – a paradigm shift in training for severe stroke. *Stroke* 2018; **49:** e239–e242.
21 Breen D, O'Brien S, McCarthy N, Gallagher A, Walshe N. Effect of a proficiency-based progression simulation programme on clinical communication for the deteriorating patient: a randomised controlled trial. *BMJ Open* 2019; **9:** e025992.
22 Angelo RL, Ryu RK, Pedowitz RA *et al.* A proficiency-based progression training curriculum coupled with a model simulator results in the acquisition of a superior arthroscopic Bankart skill set. *Arthroscopy* 2015; **31:** 1854–71.
23 Van Sickle K, Ritter EM, Baghai M *et al.* Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg* 2008; **207:** 560–8.
24 Ahlberg G, Enochsson L, Gallagher AG *et al.* Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 2007; **193:** 797–804.
25 Seymour NE, Gallagher AG, Roman SA *et al.* Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002; **236:** 458–63; discussion 463–4.
26 Gallagher AG, O'Sullivan GC, Leonard G, Bunting BP, McGlade KJ. Objective structured assessment of technical skills and checklist scales reliability compared for high stakes assessments. *ANZ J Surg* 2014; **84:** 568–73.

27 Gallagher AG, Ryu RK, Pedowitz RA, Henn P, Angelo RL. Inter-rater reliability for metrics scored in a binary fashion – performance assessment for an arthroscopic Bankart repair. *Arthroscopy* 2018; **34:** 2191–8.
28 Satava RM, Stefanidis D, Levy JS *et al.* Proving the effectiveness of the fundamentals of robotic surgery (FRS) skills curriculum: a single-blinded, multispecialty, multi-institutional randomized control trial. *Ann Surg* 2019; **272:** 384–92.
29 The Standards for Educational and Psychological Testing. *American Educational Research Association, American Psychological Association, and National Council on Measurement in Education).* Washington D.C.: American Educational Research Association, 1999. https://www.amazon.co.uk/Standards-Educational-Psychological-American-Association/dp/0935302255/ref=sr_1_2?dchild=1&qid=1596394537&refinements=p_27%3AAmerican+Educational+Research+Association&s=books&sr=1-2&text=American+Educational+Research+Association (accessed June 2020).
30 Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surgl Endosc* 2003; **17:** 1525–9.
31 Cuschieri A, Francis N, Crosby J, Hanna GB. What do master surgeons think of surgical competence and revalidation? *Am J Surg* 2001; **182:** 110–6.
32 Coleman M, Hanna G, Kennedy R, National Training Programme Lapco. The National Training Programme for laparoscopic colorectal surgery in England: a new training paradigm. *Colorectal Dis* 2011; **13:** 614–6.
33 Mehta T, Allison DB. How much variation in outcomes is too much in a center of excellence for bariatric surgery? *JAMA* 2018; **319:** 1932–3.
34 Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993; **100:** 363–406.
35 Gallagher A, Seymour NE, Jordan-Black JA, Bunting BP, McGlade K, Satava RM. Prospective, randomized assessment of transfer of training (ToT) and transfer effectiveness ratio (TER) of virtual reality simulation training for laparoscopic skill acquisition. *Ann Surg* 2013; **257:** 1025–31.
36 Gallagher AG, McClure N, McGuigan J, Crothers I, Browning J. Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR). *Endoscopy* 1999; **31:** 310–3.
37 Martin JA, Regehr G, Reznick R *et al.* Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997; **84:** 273–8.
38 Goh AC, Goldfarb DW, Sander JC, Miles BJ, Dunkin BJ. Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. *J Urol* 2012; **187:** 247–52.
39 Currie J, Bond RR, McCullagh P *et al.* Wearable technology-based metrics for predicting operator performance during cardiac catheterisation. *Int J Comput Asssit Radiol Surg* 2019; **14:** 645–57.
40 Cates CU, Gallagher AG. The future of simulation technologies for complex cardiovascular procedures. *Eur Heart J* 2012; **33:** 2127–34.
41 Gallagher AG, Cates CU. Virtual reality training for the operating room and cardiac catheterisation laboratory. *Lancet* 2004; **364:** 1538–40.
42 Pedowitz RA, Nicandri GT, Angelo RL, Ryu RK, Gallagher AG. Objective assessment of knot-tying proficiency with the fundamentals of arthroscopic surgery training program workstation and knot tester. *Arthroscopy* 2015; **31:** 1872–9.
43 Miskovic D, Ahmed J, Bissett-Amess R *et al.* European consensus on the standardisation of robotic total mesorectal excision for rectal cancer. *Colorectal Dis* 2019; **21:** 270–6.
44 Optimal surgical technique for rectal cancer. National Institute for Health and Care Excellence for colorectal cancer (update). https://www.nice.org.uk/guidance/ng151/documents/evidence-review-5 (Accessed March 29, 2020).
45 Wright JP, Albert MR. A current review of robotic colorectal surgery. *Ann Laparosc Endosc Surg* 2020; **5:** 9.
46 Larsen SG, Pfeffer F, Korner H, Norwegian Colorectal Cancer Group. Norwegian moratorium on transanal total mesorectal excision. *Br J Surg* 2019; **106:** 1120–1.
47 Deijen CL, Velthuis S, Tsai A *et al.* COLOR III: a multicentre randomised clinical trial comparing transanal TME versus laparoscopic TME for mid and low rectal cancer. *Surg Endosc* 2016; **30:** 3210–5.
48 Rouanet P, Gourgou S, Gogenur I *et al.* Rectal Surgery Evaluation Trial: protocol for a parallel cohort trial of outcomes using surgical techniques for total mesorectal excision with low anterior resection in high-risk rectal cancer patients. *Colorectal Dis* 2019; **21:** 516–22.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Sequence of the procedure assessed using the metrics in the robotic-assisted low anterior resection.

