Learning Curve Analysis for Intracorporeal Robot-assisted Radical Cystectomy: Results from the EAU Robotic Urology Section Scientific Working Group


*Department of Urology, Rijnstate Hospital, Arnhem, The Netherlands; **Department of Operating Rooms, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands; †Department for Health Evidence, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands; ‡Department of Urology, St. George’s University Hospitals, London, UK; §Department of Urology, University Hospital Gent, Gent, Belgium; ‡Department of Urology, Mount Sinai Hospital, New York, NY, USA; ‡Department of Urology, Karolinska University Hospital, Stockholm, Sweden; ‡Department of Urology, University College London Hospitals, London, UK; †Department of Urology, Onze Lieve Vrouw Hospital, Aalst, Belgium; †Department of Urology, Maassstad Hospital, Rotterdam, The Netherlands; ‡Department of Urology, Vincenz Hospital, Hannover, Germany; ‡Department of Urology, ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy; †Department of Urology, Division of Oncology, Urological Research Institute, IRCCS Ospedale San Raffaele, Milan, Italy; ‡Department of Urology, Radboud University Medical Center, Nijmegen, The Netherlands

Article info

Article history:
Accepted March 5, 2022

Keywords:
Learning curve
Radical cystectomy
Robot-assisted
Bladder cancer
Complications
Intracorporeal

Abstract

Background: The utilisation of robot-assisted radical cystectomy with intracorporeal reconstruction (iRARC) has increased in recent years. Little is known about the length of the learning curve (LC) for this procedure.

Objective: To study the length of the LC for iRARC in terms of 90-d major complications (MC90; Clavien-Dindo grade ≥3), 90-d overall complications (OC90, Clavien-Dindo grades 1–5), operating time (OT), estimated blood loss (EBL), and length of hospital stay (LOS).

Design, setting, and participants: This was a retrospective analysis of all consecutive iRARC cases from nine European high-volume hospitals with ≥100 cases. All patients had bladder cancer for which iRARC was performed, with an ileal conduit or neobladder as the urinary diversion.

Outcome measurements and statistical analysis: Outcome parameters used as a proxy for LC length were the number of consecutive cases needed to reach a plateau level in two-piece mixed-effects models for MC90, OC90, OT, EBL, and LOS.

Results and limitations: A total of 2186 patients undergoing iRARC between 2003 and 2018 were included. The plateau levels for MC90 and OC90 were reached after 137 cases (95% confidence interval [CI] 80–193) and 97 cases (95% CI 41–154), respectively. The mean MC90 rate at the plateau was 14% (95% CI 7–21%).
1. Introduction

Radical cystectomy (RC) with pelvic lymphadenectomy and urinary diversion (UD) is the standard of care for both high-risk non–muscle-invasive and muscle-invasive bladder cancer [1]. Surgical volumes (RCs per year) per centre and per surgeon appear to be related to a reduction in complications. For this procedure the European Association of Urology guidelines recommend that hospitals should perform at least ten, and preferably more than 20 RCs annually [2,3]. Over the past decade, robot-assisted RC (RARC) has been increasingly performed, with potential advantages in terms of lower blood loss without compromising oncological outcomes [4,5]. The randomised controlled trials comparing RARC with open RC (ORC) that have been published so far used the extracorporeal RARC (eRARC) technique [5–9].

The International Robotic Cystectomy Consortium has reported outcomes for 1094 totally intracorporeal RARC (iRARC) procedures over a 10-yr period. iRARC use increased from 9% in 2005 to 97% in 2016. The consortium found that Clavien-Dindo grade 3–5 (major) complications in iRARC decreased significantly from 25% in 2005 to 6% in 2015, while a similar decrease in major complications was not identified (13% in 2006 and 14% in 2015) in the eRARC cohort [10,11]. The consortium also demonstrated shorter operating time (OT) for iRARC than for eRARC, as well as lower estimated blood loss (EBL). The intracorporeal technique may offer more benefits for the patient, but it is more challenging than eRARC and may require a longer period of training to achieve good outcomes [12].

The implementation of new surgical techniques requires careful stepwise progression according to the IDEAL (Initiating, Diagnosing, Establishing, Acting, Learning) model of technical development and evaluation in order to protect patients as much as possible against potential harms associated with such implementation [13]. Assessment of the learning curve (LC) is important to understand possible difficulties when implementing and evaluating a new technique. Most studies on LC analyses have been single-centre small case series and focused mainly on a single intraoperative parameter such as mean operating time (OT), estimated blood loss (EBL), or lymph node yield (LNy). The LC range was 30–50 cases according to different definitions and LC phases [14–16]. However, multicentre LC studies for iRARC are lacking. The objective of the current study was to describe the LC for iRARC among nine European high-volume centres, all with at least 100 consecutive iRARC cases, in terms of 90-d major complications (MC90, Clavien-Dindo grade ≥3), 90-d overall complications (OC90, Clavien-Dindo grades 1–5), OT, EBL, and hospital length of stay (LOS).

2. Patients and methods

2.1. Design

We performed an analysis of prospectively maintained multicentre and multinational cohort data from nine European high-volume centres for bladder cancer surgery performing iRARC, with at least 100 cases performed and a caseload of ≥20 RCs per year. The study was designed in cooperation with the EAU Robotic Urology Section Scientific Working Group.

2.2. Patients

All consecutive patients with bladder cancer undergoing iRARC with an ileal conduit or neobladder for UD since implementation of iRARC in the different centres (between 2003 and 2018) were included. Patients were excluded if a ureterocutaneostomy or catheterisable pouch was used for UD. Baseline characteristics included sex, age, body mass index (BMI), American Society of Anesthesiologists (ASA) performance status score, clinical tumour stage, and neoadjuvant chemotherapy (NAC) status.

2.3. Outcome measurements

Surgical outcomes included mean MC90 and OC90 rates, OT, EBL, and LOS. Complications were registered according to the standardised methodology in the EAU guidelines (Supplementary Table 1) [17]. The primary aim was to investigate how many consecutive cases were needed to reach a plateau for MC90, indicating a stabilisation in the outcomes which may be considered as the completion of the LC. The secondary aim was to investigate how many consecutive cases were needed to reach a plateau for OC90, OT, EBL, and LOS.

2.4. Statistical analysis

Generalised additive mixed models (GAMMs) were used to fit the MC90, OC90, OT, EBL and LOS data [18]. In addition, a two-piece mixed-effects model was used to describe the LC using study centre as a random effect [19]. The first part is the linear descending/ascending section, which represents the learning phase, and the second part starts when the data plateau level was reached after 75 cases (95% CI 65–86) for OT, 88 cases (95% CI 70–106) for EBL, and 198 cases (95% CI 130–266) for LOS. A major limitation of the study is the difference in the balance of urinary diversion types between centres.

Conclusions: This multicentre retrospective analysis for the iRARC LC among nine European centres showed that 137 consecutive cases were needed to reach a stable MC90 rate.

Patient summary: We carried out a multicentre analysis of the surgical learning curve for robot-assisted removal of the bladder and bladder reconstruction in patients with bladder cancer. We found that 137 consecutive cases were needed to reach a stable rate of serious complications.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of European Association of Urology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
were best fitted to a horizontal line, which represents stable incidence of the outcome parameter: the plateau phase. The breaking point was the case number at the transition from the ascending/descending section to the horizontal line in the two-piece model, which represented the length of the LC. Breaking points and plateau levels were estimated along with 95% confidence intervals (CIs). Pooling of data from centres with case series of different lengths is associated with more uncertainty for results at the end of the LC. To ensure that the LC estimations were based on at least three of the nine centres, a maximum of the first 274 consecutive cases per centre was used.

To assess to what extent outcomes were related to case mix (age, sex, BMI, ASA score, NAC, and cT stage), GAMM curves for expected outcomes based on the case mix were plotted. Missing data were estimated using multiple imputation with the creation of 25 data sets. Using the imputed data set, a prediction model was developed. The predicted outcomes for multiple imputation with the creation of 25 data sets. Using the imputed data set, a prediction model was developed. The predicted outcomes for the data sets created were pooled, plotted, and visually compared to the data set, a prediction model was developed. The predicted outcomes for the data sets created were pooled, plotted, and visually compared to the outcome parameters observed in order to assess whether it was likely that changes in outcomes could be explained by trends in patient selection [20]. Data analyses were performed using R version 4.0.4 (R Foundation for Statistical Computing, Vienna, Austria).

### 3. Results

#### 3.1. Patients

Between 2003 and 2018, a total of 2186 patients underwent iRARC in nine centres. The mean number of iRARC cases per year per centre was 33 (range 17–74, interquartile range 22.4–30.4). In most of the cases (n = 1658) an ileal conduit was used for UD, versus 528 neobladders. The mean age was 68 yr and the mean BMI was 26.5 kg/m². NAC was administered in 24% of the patients. The minimum follow-up was 90 d. All the patient characteristics are presented in Table 1.

#### 3.2. Outcomes

The number of cases needed to reach the breaking point towards a plateau in the mixed-effects model (and representing the length of the LC) for MC90 was 137 cases (95% CI 80–193). A plateau was reached at 14% (95% CI 7–21). The length of the LC was 97 cases (95% CI 41–154), for OC90, 75 cases (95% CI 65–86) for OT, 88 cases (95% CI 70–106) for EBL, and 198 cases (95% CI 130–266) for LOS.

The plateau was 48% (95% CI 38–57) for OC90, 321 min (95% CI 293–349) for OT, 292 ml (95% CI 230–353) for EBL, and 9.5 d (95% CI 7.7–11.3) for LOS. The GAMM curves (orange lines) and the two-piece LCs (black lines) are depicted in Fig. 1. In addition, predicted outcomes based on the case mix were stable over time (dashed grey lines; Fig. 1). Details on the complications reported by Clavien-Dindo grade can be found in Supplementary Table 2.

### 4. Discussion

This large European retrospective cohort study confirms that the LC for iRARC is associated with a large number of consecutive cases to reach a plateau in the two-piece model. For all outcomes measured, this number was substantially greater than the 30–50 cases previously reported: 137 cases (95% CI 80–193) for MC90; 97 cases (95% CI 41–154) for OC90; 75 cases (95% CI 65–86) for OT; 88 cases (95% CI 70–106) for EBL; and 198 cases (95% CI 130–266) for LOS [14–16].

Guru et al [14] were the first to analyse the LC for eRARC in their single-centre series of 100 cases in 2009. The plateau (defined as the number of cases needed for a change of <1% for a variable) was reached after 16 cases for OT, 11 cases for EBL, 12 cases for LOS, and 30 cases for LNY. The mean plateau (after 75 cases) for OT in our iRARC series was 321 min (95% CI 293–349), which is shorter than the OT mean plateau of 343 min in the eRARC series reported by Guru et al. In addition, our plateau for EBL was lower, at a mean of 292 ml (95% CI 230–353) versus 598 ml.

Hayn et al [15] reported an “acceptable level of proficiency” by the 30th eRARC case for proxy measures of surgical quality. The definition used in their LC analysis was the number of cases needed to reach a predefined level of expertise (PLE) in terms of OT <390 min, a positive surgical margin (PSM) rate <5%, and LNY >20. The length of the consecutive case series used in the analysis varied widely between the 14 institutions (range 4–119). The authors reported that only the LC for OT reached a plateau, at a mean of 386 min. In the current study with larger case series per centre (all centres had >100 cases and three of the
nine centres had ≥274 cases), we were able to show plateau levels for all outcomes.

Dell’Oglio et al [16] studied the effect of surgical experience (consecutive case numbers for iRARC) on OT, complications within 30 d (Clavien-Dindo grade ≥2), and 18-mo recurrence rates in a single-centre case series of 164 patients. A LOWESS function revealed a nonlinear relationship between surgical experience and OT, with stable OT after 50 cases, whereas a plateau was not reached for complications or recurrences after 88 cases, indicating that the LC is longer for these outcomes that are harder to achieve.

A major strength of our work is that, to the best of our knowledge, it is the first study of a European multinational database from institutions with large iRARC series (>100 per centre) with information for all consecutive cases. Another strength is the method we used to compute the length of the LCs, taking into account the individual centres. We relied on the standardised methodology proposed by the EAU ad hoc panel on reporting and grading complications after urological surgical procedures that is proven to avoid missing critical information that could lead to underestimation of perioperative complications [17,21]. To the best of our knowledge, this study represents one of the few reports in the specific setting of RARC that fulfils all 14 criteria proposed by the EAU ad hoc panel [22].

Some limitations should also be mentioned. First, although we aimed to obtain a homogeneous group of patients by including only iRARC, local variations in surgical technique, the numbers of surgeons per centre, and balance in the type of UD between centres might have influenced the results, but we chose to take a pragmatic approach to be able to include sufficient cases to answer our research questions. Because of the more complex and time-consuming intracorporeal neobladder construction, centres may have preferentially opted for ileal conduits at the start of their iRARC programme, although we think that surgical teams and individual surgeons learn from each type of UD during iRARC. We analysed the consecutive case series for each centre, consisting of a mix of ileal conduit and neobladder UDs. Centres started their iRARC programmes in different years and the balance between ileal conduit and neobladder UDs differs among the centres. The data in this study include results from pioneer centres and early adopters of iRARC. Almost 90% of the cases were performed between 2012 and 2018. A sensitivity analysis with cases from 2012 onwards showed a breaking point for MC90 at 128 cases and a plateau at 14%, which demonstrates that
our results are robust. Details of the balance between ileal conduit and neobladder UDPs per year of surgery and per centre, as well as mean complication rates and median OT, EBL, and LOS by diversion type are summarised in Supplementary Tables 3 and 4.

Second, there may be selection bias in the case series. However, the expected outcomes plotted on the basis of case mix were very stable over time, which suggests that the improvement in outcomes could not be explained by patient selection alone (Fig. 1).

Third, this is a retrospective analysis of data from nine different centres. Different perioperative protocols may have been used among centres and over time. In addition, information about which robotic system was used is lacking. The newer devices (da Vinci X and Xi; Intuitive Surgical, Sunnyvale, CA, USA) may have advantages in performing total intracorporeal reconstructions compared to older systems from the same provider (da Vinci S and Si). Previous experience in robot-assisted radical prostatectomy (RARP) may influence the LC for RARC [15]. The surgeons in the current study all had previous experience in robot-assisted surgery and all centres performed >20 iRARC cases per year. Therefore, the present results may not be applicable to robot-naïve surgeons or to hospitals with a lower annual caseload. Details regarding the experience of the surgeons in our study can be found in Supplementary Fig. 1.

Fourth, information about LNY, PSM rates, and oncological and long-term functional outcomes is missing in our database. Therefore, we cannot indicate the length of the LC for these important outcomes. Long-term functional and oncological outcomes such as stricture rates, PSM rates, LNY, and survival are very important and should be included in future studies on LCs for iRARC.

Fifth, an enhanced recovery after surgery (ERAS) protocol can have an influence on outcomes [23]. In the time period for this analysis, perioperative protocols may have been modified and no standardised protocol was used for this study. In addition, health care systems vary by country and region and can affect LOS. In the current study, LOS continues to improve, which may reflect the improvements in OT, EBL, and complication rates observed and the introduction of ERAS protocols, which could explain the high number of consecutive cases needed to reach a breaking point for LOS.

Several observational studies suggest that iRARC offers clinical benefits for the patient compared to eRARC and ORC, such as fewer complications, less EBL, and shorter LOS [4,11,24]. In addition, one single-centre study reported a decrease in ureteroenteric anastomotic strictures from 17.5% to 4.9% after 75 consecutive cases [25]. An ongoing phase 3 prospective multicentre randomised study comparing ORC to iRARC (iROC trial, ClinicalTrials.gov NCT03049410) may provide evidence on perioperative outcomes [26].

While iRARC may be advantageous, centralisation of RC remains crucial in improving outcomes as there is a clear relation between RC volume and outcomes [2]. The current study shows that the LC for iRARC is long and that the mean number of cases needed to reach a breaking point depends on the outcome being measured. Another definition of the LC could be the number of cases needed to reach one of the three phases of learning: competency, proficiency, and mastery [27]. The first phase of the LC is generally described as improvement of operative parameters, especially OT, while most importantly focusing on patient safety. The skill level denoted as competency has frequently been used for this phase. For the second phase, a further reduction in postoperative complications and stabilisation of the OT are characteristic. Proficiency best fits the description of this second phase. When the OT and complication rate reach a plateau, mastery is achieved, which defines the third phase. Our primary outcome (MC90) reached a breaking point after 137 cases. The plateau for MC90 was 14%, which is substantially lower than the 22% reported for MC90 in the RAZOR trial, in which an extracorporeal technique was used. According to our LC analysis for iRARC, an MC90 rate of 22% would be reached after 35 cases and an MC90 rate of 20% after 64 cases. In addition, our analysis shows that outcomes continued to improve until a plateau was reached, which in our opinion indicates the length of the LC. The plateau levels for the outcomes reported in this study may be used as a reference PLE for other centres that have started or will be starting iRARC in their practice and for future training curricula.

During the learning phase, patients may be exposed to a higher risk of surgical-related morbidity. Centralisation of iRARC may lead to a higher annual case load, which may help in shortening the LC. In addition, robot-assisted surgery offers a platform for training new surgeons outside the operating room before starting (modular) training on patients [28]. In recent years, structured training programmes for the next generation of robotic surgeons have been developed and validated for RARP and robot-assisted partial nephrectomy, which have shortened the LC [29,30]. Proficiency-based progression (PBP) simulation training reduced performance error by 60% [31]. A first structured training programme for iRARC was developed with the goal of aiding surgeons to overcome the LC for this procedure, improving patient safety at the same time [32]. This iRARC training programme incorporating PBP needs to be validated.

5. Conclusions

This multicentre retrospective analysis of the iRARC LC among nine European centres with >100 consecutive cases showed that 137 consecutive cases were needed to reach a stable rate of 90-d major complications. Centralisation and structured training curricula may help to reduce the length of the LC.

Author contributions: Carl J. Wijburg had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Wijburg, Hannink, Rovers, Witjes.
Acquisition of data: Wijburg, Hannink, Michels.
Analysis and interpretation of data: Hannink, Wijburg, Rovers, Witjes.


STATISTICAL ANALYSIS: Hannink, Wijburg.

OBTAINING FUNDING: Witjes.

ADMINISTRATIVE, TECHNICAL, OR MATERIAL SUPPORT: None.

SUPERVISION: Witjes, Rovers.

OTHER (PRINCIPAL INVESTIGATORS IN PARTICIPATING CENTRES; ENROLLMENT OF STUDY PARTICIPANTS AND REGISTRATION OF CLINICAL DATA): Wijburg, Klaver.

FINANCIAL DISCLOSURES: Carl J. Wijburg certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (e.g., employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Carl J. Wijburg, Karel DeCaeuckelaer, Peter Wiklund, Abolfazl Hosseini, Frédéric D'Hondt, Sjoerd Klaver, and Sebastian Edeling are proctors for Intuitive Surgical. Ashwin Sridhar is a proctor for Intuitive Surgical and a consultant for Medtronic and CMR. Alexandre Mottrie is a consultant for Intuitive Surgical, Conmed, Medtronic, and Medicaroid, and is CEO of ORSI Academy. The remaining authors have nothing to disclose.

FUNDING/SUPPORT AND ROLE OF THE SPONSOR: This work was supported by VICI grant 91818617 from the Dutch Research Council (NWO/ZonMw). The sponsor played a role in analysis of the date.

ACKNOWLEDGMENTS: We thank Marike Ulehake for administrative support.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data to this article can be found online at https://doi.org/10.1016/j.euros.2022.03.004.

REFERENCES


