Laparoscopic radical prostatectomy outcome data: how should surgeon’s performance be reported? A retrospective learning curve analysis of two surgeons

Sarah Mason¹, Mieke Van Hemelrijck¹, Ashish Chandra², Christian Brown² and Declan Cahill², ³

¹King’s College London, Division of Cancer Studies, Cancer Epidemiology Group, London, UK
²King’s Health Partners, London SE1 9RT, UK
³Royal Marsden Hospital, London SW3 6JJ, UK

Correspondence to: Declan Cahill. Email: declan.cahill@rmh.nhs.uk

Abstract

Objective: To document the learning curve for the laparoscopic radical prostatectomy (LRP) procedure and discuss the optimal usage of prospectively documented outcome data for reporting a surgeon’s performance.

Materials and methods: Using prospectively collected data from the first series of patients to undergo LRP by two surgeons in the same institution, linear and logistic regression multivariate analyses per 25 patients were carried out to graphically represent the surgical learning curve for operative time, blood loss, complications, length of stay (LOS), and positive margins. Surgeon A carried out 275 operations between 2003–2009; Surgeon B carried out 225 between 2008–2012.

Results: Learning curves showing continuous improvement of each of the above outcomes were demonstrated for both cohorts. For surgeon A, a plateau was observed for LOS and T2 positive margins after 100 and 150 surgeries respectively. No such plateau was observed for surgeon B.

Conclusion: On documenting these learning curves and discussion of the reporting methods used, we concluded that the most informative outcome measure, with the least potential observer bias was T2 positive margins. Whether as a single measure or in combination with others, this has potential for use as an objective outcome representative of improvement in a surgeon’s skill over time.

Keywords: prostatic neoplasms, prostatectomy, learning curve, patient outcome assessment, surgeons

Published: 06/07/2016 Received: 11/12/2015

cancer 2016, 10:651 DOI: 10.3332/ecancer.2016.651

Copyright: © the authors; licensee ecancermedicalscience. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Background

Radical prostatectomy (RP) surgical outcome data is increasingly being reported in terms of a learning curve, as influenced by surgeon-specific factors [1]. There is no definitive definition though learning curves are generally expected to plateau as a surgeon masters a technique. With the potential for altering guideline recommendations, they should be interpreted and documented with caution and appropriate statistical analysis [2].

To date, learning curves in urology have demonstrated changes in outcomes—such as operative time, blood loss, LOS, positive surgical margin (PSM) rate, complications, and cancer recurrence—to be predictive of a surgeon’s experience [3–7]. Specifically for laparoscopic RP (LRP) there is variation in the number of cases estimated to be required for competency: 51 cases according to complication rates; 110 cases according to blood loss, operative time and PSM rate [3]; and 250 according to recurrence and PSM [7, 8].

With a recent push for public reporting of a surgeon’s performance within urology [9], it is hoped that increased transparency and drive for improvement will follow. The influence of the learning curve on these reports, and the extent to which this is influenced by case mix will be of paramount consideration in documenting and interpreting a surgeon’s data with application to surgical training and clinical practice. Most surgeons do not have facilities for collecting patient reported outcomes, though this is expected to change with the demand for published outcomes increasing. The current British radical prostatectomy data set relies on a surgeon’s input data even though much of which is incomplete and difficult to verify [10, 11].

Therefore, this study involved evaluation of the optimal usage of learning curves in review of a surgeon’s performance. We aimed to use LRP outcome data from two surgeons and a discussion of the outcomes used, to determine how best to report changes in a surgeon’s performance over time. Ultimately we wanted to identify whether there are surrogate measurements available that are independent of the resource (e.g. histopathology instead of surgeon-reported data).

Methods

Study cohort

We used data from the first cohort of cases of extraperitoneal LRP carried out by two surgeons at King’s Health Partners. Both surgeons performed standard laparoscopic antegrade radical prostatectomies [32]. Patients data for surgeon A (275 patients) and surgeon B (225 patients) were retrieved from electronic patient records that had been recorded prospectively. Patient consent was not required as data collection was part of the clinical audit. Surgeon A operated between April 2003 and June 2009 and surgeon B between October 2008 and October 2012. The observation period and frequency of surgeries was different between the two surgeons as this reflects the uptake of radical prostatectomy in the UK with time. Both surgeons were new consultants with no independent experience so that our observations reflect a true learning curve. They were both UK trainees and completed their training with dedicated laparoscopic fellowships. Pre-operative variables included age, prostate-specific antigen (PSA), tumour stage, and Gleason score.

Outcomes

Surgeon-dependent variables of interest were classified as operative or postoperative. Operative variables included: operative time, blood loss, and complications. Complications were reported by the operating surgeon and subsequently classified according to the 2004 Clavien-Dindo Classification [12]. Postoperative variables included: complications, LOS, and surgical margin status. The latter was assessed by three consultant histopathologists with a specialist interest in urological pathology. Prior to 2011, partial embedding of the prostate in small cassettes was the routine practice. Since 2011, the lab was progressively able to support complete embedding of the prostate in megablocks [33].

Finally, we included information on surgical stage and Gleason score. Gleason scores were divided into three categories (≤ 6, 7, and ≥ 8) as has been done in similar studies [13]. Clinical and surgical stages were classed as T0, T1a/b, T1c, T2, T3, or T4 to allow clear distinctions in analyses between patients with organ-confined and non-organ confined disease.

Apart from histopathological information, all information was prospectively collected by the surgeons.
Statistical analysis

For each cohort of patients the following analyses were carried out:

Patients were assigned a number according to ascending date of surgery as a representative value for a surgeon’s experience. For initial analysis of the effects of a surgeon’s experience on surgical outcomes, comparison was made between consecutive groups of 25 patients. Age, pre-operative PSA, Gleason category, and clinical stage were included in the multivariable regression model. All p-values were two-sided and considered statistically significant < 0.05. Linear regression models were used to determine the association between consecutive group numbers and mean operative time, blood loss, and LOS. These associations were graphically represented using cubic splines. Using logistic regression, we determined the association between consecutive group number and presence or absence of complications (preoperative and postoperative), positive margins, T2 positive margins, and T3/T4 positive margins, all graphically represented.

Since there was no change in practice over the study period, stage migration and date of surgery did not need to be accounted for in these multivariate models.

Results

Cohort details

Baseline characteristics of the patients in cohort A and B and their intra- and postoperative data are shown in Table 1. Median age was 61 (IQR 56–65) and 62 (57–66) respectively, and median presenting PSA 7 (5.6–10) and 8.3 (6.1–12.5). In both cohorts, T2 was the most common clinical stage (45% and 60% respectively) and this increased to 65% and 68% in each cohort at surgical staging. Mean operative time for surgeon A was 159 mins and for surgeon B was 225 mins; mean estimated blood loss was 268 mL and 339 mL respectively. For both surgeons mean length of a patient’s hospital stay was two days. For each outcome, there was a downward trend of continuous improvement as number of surgeries increased.

Operative time, blood loss, and length of stay

Figure 1 displays the decrease in operative time, blood loss, and LOS observed over time. The associated crude linear regression models showed that per 25 extra surgeries, mean operative time decreased by 7.5 mins for surgeon A (p < 0.001) and 10.3 mins for surgeon B (p < 0.001). A similarly shaped trend for both surgeons is seen in Figure 1. Likewise, the crude linear regression models for blood loss demonstrated a statistically significant decrease per 25 extra surgeries of 14.1 mL (p < 0.001) for surgeon A and 19.6 mL (p < 0.001) for surgeon B. The crude linear regression model for LOS showed a decrease of 0.3 nights (p < 0.001) for surgeon A and 0.1 nights (p < 0.001) for surgeon B per 25 extra surgeries. As demonstrated in Figure 1, mean LOS per 25 surgeries for surgeon A reached a plateau after 100 operations. Additional adjustments for potential confounders did not alter the above findings (results not shown).

Surgical complications

Increasing surgeon experience was associated with a decrease in surgical complications (p: 0.015 and 0.095 for surgeon A and B respectively). For surgeon A, the crude logistic regression model showed that the risk of complications decreased by 19% per 25 extra surgeries (95% CI: 0.73–0.91), while for surgeon B the risk decreased by 6% per 25 extra surgeries (95% CI: 0.81–1.08). Adjustment for potentially confounding variables did not alter these findings (results not shown). On analysis of the distribution of Clavien-Dindo scores in each group of 25 patients, there was a statistically significant difference for the types of postoperative complications in both surgeon A and surgeon B data (p: 0.040 and 0.021 respectively), but not the types of intra-operative complications (p: 0.325 and 0.530 respectively).

Positive margins

Figures 2a and b report the trend for total positive margins as number of performed surgeries increases, stratified according to surgical stage. The logistic regression models showed a decrease in total positive margins by 2% per 25 extra surgeries for surgeon A, whereas this
was 9% for surgeon B—though these associations were not statistically significant. Amongst the patients with T2 disease in surgeon A’s cohort, a statistically significant association was found between PSM rate and number of surgeries performed after adjustment for Gleason, age, PSA, and clinical stage: 29% decrease per 25 surgeries (95% CI: 0.56–0.90). A decrease of 12% (95% CI: 0.80–1.05) was reported for surgeon B. Surgeon A’s T2 positive margin rates resulted in a plateau after 150 cases (Figure 2a) whereas PSM rates were more consistent for surgeon B (Figure 2b). Such a pattern was less clear for the subgroup with T3/T4 disease.

Table 1. Baseline patient characteristics.

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>Surgeon A</th>
<th>Surgeon B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age at surgery; years (IQR)</td>
<td>61 (56–65)</td>
<td>62 (57–66)</td>
</tr>
<tr>
<td>Number of operations per year (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>29 (10.5)</td>
<td>-</td>
</tr>
<tr>
<td>2004</td>
<td>31 (11)</td>
<td>-</td>
</tr>
<tr>
<td>2005</td>
<td>30 (11)</td>
<td>-</td>
</tr>
<tr>
<td>2006</td>
<td>48 (17)</td>
<td>-</td>
</tr>
<tr>
<td>2007</td>
<td>61 (22)</td>
<td>-</td>
</tr>
<tr>
<td>2008</td>
<td>49 (18)</td>
<td>9 (4)</td>
</tr>
<tr>
<td>2009</td>
<td>29 (10.5)</td>
<td>56 (25)</td>
</tr>
<tr>
<td>2010</td>
<td>-</td>
<td>66 (29)</td>
</tr>
<tr>
<td>2011</td>
<td>-</td>
<td>60 (27)</td>
</tr>
<tr>
<td>2012</td>
<td>-</td>
<td>34 (15)</td>
</tr>
<tr>
<td>Median presenting PSA; ng/mL (IQR)</td>
<td>7.0 (5.6–10.0)*</td>
<td>8.3 (6.1–12.5)</td>
</tr>
<tr>
<td>Clinical stage; number of patients (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1a/b</td>
<td>11 (4)</td>
<td>8 (3.5)</td>
</tr>
<tr>
<td>T1c</td>
<td>122 (45)</td>
<td>71 (31.5)</td>
</tr>
<tr>
<td>T2</td>
<td>123 (45)</td>
<td>135 (60.0)</td>
</tr>
<tr>
<td>T3</td>
<td>16 (6)</td>
<td>11 (5.0)</td>
</tr>
<tr>
<td>Gleason; number of patients (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤6</td>
<td>140 (52)</td>
<td>88 (39)</td>
</tr>
<tr>
<td>7</td>
<td>104 (39)</td>
<td>103 (46)</td>
</tr>
<tr>
<td>8–10</td>
<td>23 (9)</td>
<td>34 (15)</td>
</tr>
<tr>
<td>Intra- and post-operative characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean time in min (SD)</td>
<td>159 (37)**</td>
<td>225 (58)</td>
</tr>
<tr>
<td>Mean blood loss in mL (SD)</td>
<td>268 (212)**</td>
<td>339 (230)</td>
</tr>
<tr>
<td>Intra-operative complications; number of patients (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>19 (7)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>No</td>
<td>256 (93)</td>
<td>223 (99)</td>
</tr>
<tr>
<td>Clavien-Dindo; number (% of total intraoperative complications)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade I</td>
<td>9 (47.0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Grade II</td>
<td>2 (10.5)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Grade III</td>
<td>6 (32.0)</td>
<td>2 (100)</td>
</tr>
<tr>
<td>Grade IV</td>
<td>2 (10.5)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Postoperative complications; number of patients (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>39 (14)</td>
<td>29 (13)</td>
</tr>
<tr>
<td>No</td>
<td>236 (86)</td>
<td>196 (87)</td>
</tr>
<tr>
<td>Clavien-Dindo; number (% of total postoperative complications)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade I</td>
<td>13 (33)</td>
<td>12 (41.5)</td>
</tr>
<tr>
<td>Grade II</td>
<td>13 (33)</td>
<td>5 (17.0)</td>
</tr>
<tr>
<td>Grade III</td>
<td>8 (21)</td>
<td>12 (41.5)</td>
</tr>
<tr>
<td>Grade IV</td>
<td>5 (13)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>
Baseline patient characteristics and intra-/postoperative patient characteristics for total cohort of patients operated on by Surgeon A and Surgeon B.

IQR: interquartile range; SD: standard deviation; *Data missing for 4 patients; **Data missing for 11 patients; + Data missing for 52 patients; ++ Data missing for 1 patient

### Discussion

Our study shows that operative time, blood loss, complications, hospital stay, and T2 margin status illustrate continuous improvement as a surgeon’s experience increases. We observed a plateau in LOS in hospital and rate of T2 PSMs.

The trends in operative time, blood loss, LOS, and complications described here parallel those which have previously been published [14, 15]. Despite some conflicting results, a recent review concluded PSM rate for open, perineal, and laparoscopic prostatectomy procedures improved with a surgeon’s skill across the learning curve [1]. Furthermore, PSM rates have been shown to be lower amongst more experienced higher volume surgeons [16, 17].

Concerning the paucity of data on how a surgical learning curve should be documented in the current context of an upcoming surgeon’s performance publication [18], it is necessary to take into account the way in which surgical outcomes are reported in an attempt to evaluate the optimal reporting method. For example, surgeon documented blood loss relies on volume estimations in fluids mixed with urine, saline, and wash, and the bleeding complications involve a method and cost-dependent balance between tissue sparing and bleeding. Furthermore, as acknowledged in the British Association of Urological Surgeons nephrectomy database, outcomes documented by operating surgeons may be subject to bias, relying on honest, complete, and consistent reporting [19]. While Hospital Episode Statistics data has advantages and can be used to verify data such as LOS, it too may be subject to certain inaccuracies when documented [20]. Within our own cohort, there was notable variation within blood loss data reported by each individual surgeon from case to case, as well as significant variation between the surgeons in reporting operative time (data not shown). Given this variation, the clinical relevance and utility of reporting changes per patient is questionable. Aside from variations in surgeon reporting, LOS also demonstrates the possible impact of other factors such as the day of the week on which a surgeon operates or the effects of other health care professionals involved in a patient’s care following surgery [15], [21].

In contrast to the above, PSM status is an outcome acquired independently of surgeon self-reporting, sensitive to continuous monitoring of a surgeon’s skill [22]. There is little evidence of significant inter-observer variability between pathologists [23]. A T2 positive margin is likely a result of iatrogenic capsular incision, being much less stage-dependent than non-organ confined cases. T3 prostate cancers are a very heterogeneous group. T3 prostate cancer positive margins are more tumour dependent than surgeon dependent. The opposite is the case with T2 positive margins. Regardless of the volume of cancer within the prostate, if the gland is not incised and the cancer is organ confined, there will be a negative margin. Though there is conflicting evidence as to whether or not surgical margins status is a useful indicator of prognosis [8], [13], [24–27], we propose T2 margin rate to be most useful as described here. It is a marker of a surgeon’s performance, improving with a surgeon’s experience alongside various other surgical parameters.
One of the strengths of this study is the division of cases into groups of 25 operations—in a similar manner to the plotting of case number [28] as opposed to groups of at least 50 [1]—thereby decreasing the risk of drawing misrepresentative conclusions regarding the number of cases required for competency [29]. Multivariate analysis accounted for the positive association found in these cohorts between a surgeon’s experience and Gleason score and clinical stage. Heterogeneity between surgeons is highlighted by learning curve studies such as this, demonstrating the need for individual analysis of a surgeon’s performance with implications for both practice and training [30]. While a follow-up study might be proposed which includes the five-year biochemical recurrence (BCR) rates and continence data, it is beneficial that this study has used outcomes immediately available for a continuous record of performance. This would also be facilitated by the use of new early markers such as ultrasensitive PSA [31].

Lack of standardised measures of continence and potency impedes analysis of functional outcomes [3] and retrospectively derived data can prove unreliable. Although the surgeons in this study underwent training in a different environment to current trainees, the same outcome measures are useful and relevant for current practice. Clearly documented learning curves are likely to have future implication for the training and equipping of inexperienced surgeons in order to minimise the learning curve in surgical practice. Finally, it is a limitation that we only compared two surgeons—and future studies comparing more surgeons will provide even more insight into the use of T2 positive margin status as a marker for a surgeon’s experience.

**Conclusion**

Given the need for clarity and transparency of a surgeon’s proficiency in reporting of surgeon-specific outcomes and learning curves, this study was undertaken. It shows that T2 positive margin status is the best objectively acquired parameter representative of a surgeon’s proficiency which improves with experience. All outcomes have their importance especially in terms of long-term disease-free survival and functional well being. Until outcome analysis becomes a routine part of a clinical practice, dependent on staffing and funding, it would helpful to have a surrogate to gauge performance which is least subject to reporting bias.
Figure 2. Positive margin rates: (a) positive margin rate per 25 consecutive operations (total, T2 and T3/T4 shown in blue, red, and green respectively) for surgeon A; (b) positive margin rate per 25 consecutive operations (total, T2 and T3/T4 shown in blue, red, and green respectively) for surgeon B.

Conflict of interest and funding

The authors declare that they have no conflict of interest. There are no declarations of financial support. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors’ contribution

S Mason: Literature research, data analysis, manuscript writing.
M Van Hemelrijck: Project proposal, data analysis, manuscript editing.
D Cahill: Project development, data collection, manuscript editing.
C Brown: Data collection, manuscript editing.
A Chandra: Data collection, manuscript editing.

All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgment

This research was supported by the Experimental Cancer Medicine Centre at King’s College London and also by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Ethical approval: The manuscript does not contain patient data and consent was not required as data collection was part of the clinical audit.
Guarantor

SM

References

9. Iacobucci G (2012) Performance data on all surgeons in England will be published within two years Brit Med J 345 e8377 DOI: 10.1136/bmj.e8377 PMID: 23220598


