Assessment of Quality Outcomes for Robotic Pancreaticoduodenectomy
Identification of the Learning Curve

Brian A. Boone, MD; Mazen Zenati, MD, PhD; Melissa E. Hogg, MD; Jennifer Steve, BA; Arthur James Moser, MD; David L. Bartlett, MD; Herbert J. Zeh, MD; Amer H. Zureikat, MD

IMPORTANCE Quality assessment is an important instrument to ensure optimal surgical outcomes, particularly during the adoption of new surgical technology. The use of the robotic platform for complex pancreatic resections, such as the pancreaticoduodenectomy, requires close monitoring of outcomes during its implementation phase to ensure patient safety is maintained and the learning curve identified.

OBJECTIVE To report the results of a quality analysis and learning curve during the implementation of robotic pancreaticoduodenectomy (RPD).

DESIGN, SETTING, AND PARTICIPANTS A retrospective review of a prospectively maintained database of 200 consecutive patients who underwent RPD in a large academic center from October 3, 2008, through March 1, 2014, was evaluated for important metrics of quality. Patients were analyzed in groups of 20 to minimize demographic differences and optimize the ability to detect statistically meaningful changes in performance.

EXPOSURES Robotic pancreaticoduodenectomy.

MAIN OUTCOMES AND MEASURES Optimization of perioperative outcome parameters.

RESULTS No statistical differences in mortality rates or major morbidity were noted during the study. Statistical improvements in estimated blood loss and conversions to open surgery occurred after 20 cases (600 mL vs 250 mL [P = .002] and 35.0% vs 3.3% [P < .001], respectively), incidence of pancreatic fistula after 40 cases (27.5% vs 14.4%; P = .04), and operative time after 80 cases (581 minutes vs 417 minutes [P < .001]). Complication rates, lengths of stay, and readmission rates showed continuous improvement that did not reach statistical significance. Outcomes for the last 120 cases (representing optimized metrics beyond the learning curve) included a mean operative time of 417 minutes, median estimated blood loss of 250 mL, a conversion rate of 3.3%, 90-day mortality of 3.3%, a clinically significant (grade B/C) pancreatic fistula rate of 6.9%, and a median length of stay of 9 days.

CONCLUSIONS AND RELEVANCE Continuous assessment of quality metrics allows for safe implementation of RPD. We identified several inflexion points corresponding to optimization of performance metrics for RPD that can be used as benchmarks for surgeons who are adopting this technology.

Published online March 11, 2015.
In recent years, a focus on quality outcome measures has dominated the national discourse on health care delivery and funding. Health care is in the midst of a paradigm shift, with reimbursement moving from fees for service to payment for performance, which sets rates based on patient quality outcomes and satisfaction. This strategy is designed to increase competition among hospitals and physicians and to lower costs by rewarding the most effective health care professionals. Outcome data are being distributed to regional hospitals and physicians to compare performance measures. Early results from this heightened awareness of outcome measures have been encouraging, suggesting that the movement toward identifying optimal performance metrics and reimbursement based on quality will continue to gain momentum.

This environment raises important questions regarding how to monitor quality and patient safety during implementation of new surgical technology and innovative techniques, such as the use of robotics in surgery. Robotic surgery is becoming increasingly used in complex operations, including pancreatic resections. However, despite this small number of studies demonstrating the feasibility of this approach for pancreaticoduodenectomy (PD), no study, to our knowledge, has attempted to review the outcomes and learning curve of robotic PD (RPD) during its implementation phase. Such an analysis is important for surgeons who are adopting this platform and establishes a benchmark for which comparisons of RPD to the open or laparoscopic approach can be performed.

We describe our experience during implementation of the robotic platform for the first 200 consecutive RPDs. By monitoring outcomes in groups of 20 patients, we identified the number of cases necessary to reach optimal performance for a number of important metrics while ensuring that safety and quality are maintained during this initial experience.

## Methods

### Patient Selection

After obtaining institutional review board approval from the University of Pittsburgh, a retrospective review of a prospectively maintained database of all RPDs from October 2008 through March 2014 was conducted. This cohort represents our first 200 consecutive RPDs. Patients with planned major vascular resections (end-to-end primary anastomoses or use of harvested vein graft) were considered ineligible for robotic resection during that period.

### Quality Outcome Analysis

Patient demographics and perioperative outcomes were evaluated to detect variations in outcomes and monitor for selection bias. Outcomes were reviewed in groups of 20 cases to minimize the effects of demographic differences and optimize normalcy in distribution. Outcomes were plotted on control charts, allowing for visualization of outliers in each cohort as well as throughout the study period. Groups harboring variation were identified, and independent case reviews were undertaken at regularly scheduled meetings with surgeons, a statistician, and research staff (B.A.B., M.Z., M.E.H., J.S., A.J.M., D.L.B., H.J.Z., and A.H.Z) to identify the root causes of the variability. All outcomes were followed up for 90 days.

The operative time reported was from skin incision until closure. Drain fluid amylase levels were evaluated on or after postoperative day 3, and a value greater than 3 times the serum amylase level was considered a pancreatic fistula. The pancreatic fistulas were graded (A, B, or C) according to the International Study Group on Pancreatic Fistula (ISGPF) classification. Perioperative morbidity was scored based on the Clavien-Dindo classification.

### Statistical Analysis

Continuous data that exhibited normalcy were summarized in means and SDs. Data that were not normally distributed were summarized in medians and interquartile ranges. Categorical data were summarized as frequencies and percentages. Analyses of variance and 2-tailed unpaired t test were used to analyze continuous data that were normally distributed; otherwise, the related nonparametric Kruskal-Wallis and Wilcoxon rank sum tests were used. Frequency data were tested using Fisher exact test. Adjustment for patient characteristics and subgroup analysis was performed. Cumulative sum analysis, which graphically depicts the cumulative sum of differences from the total cohort’s mean, was used to depict optimization of performance throughout the study. \( P < .05 \) was considered statistically significant.

## Results

### Patient Demographics and Outcomes

Patient demographics were evaluated for 200 consecutive RPDs to monitor for selection bias that might influence the outcomes of a particular group (Table 1). The RPDs were performed to treat 82 pancreatic ductal adenocarcinomas (41.0%), 17 neuroendocrine tumors (8.5%), 60 cases of periampullary cancer (30.0%), 25 indeterminate cystic lesions (12.5%), 12 benign lesions (6.0%), 2 gastrointestinal stromal tumors (1.0%), and 2 cases of metastatic disease (1.0%). One subset of patients, cases 41 through 60, had a lower percentage of malignant disease compared with the rest of the patients (50.0% vs 83.0%; \( P = .02 \)). In addition, more neoadjuvant therapy was used in the latter part of the study period. The 90-day mortality for the entire period was 3.3%. The mean (SD) operative time was 483 (113) minutes, with median estimated blood loss of 250 mL (interquartile range, 150-500 mL). Conversion to open surgery was required for 13 patients (6.5%). Thirty-four patients (17.0%) developed a pancreatic fistula (ISGPF grade A, 17 [50.0%]; B, 10 [29.4%]; or C, 7 [20.6%]). A total of 135 patients (67.5%) experienced a complication, with major complications (Clavien-Dindo classification ≥3) occurring in 52 (26.0%). The median length of stay was 9 days, with a 90-day readmission rate of 29.2%. An R0 resection was performed for 184 patients (92.0%), with a median lymph node harvest of 22.
**Optimization of Safety Outcomes**

Patient safety outcomes included estimated blood loss, pancreatic fistula rate, and complications. Estimated blood loss showed a significant improvement after 20 procedures (600 mL vs 250 mL; \( P < .002 \)) (Figure 1A). The overall pancreatic fistula rate (sum of all ISGPF grades) improved after 40 cases (27.5% vs 14.4%; \( P = .04 \)) (Figure 1B). The rate for clinically significant pancreatic fistula (ISGPF grade B/C) improved from 15.0% (6 patients) in the first 40 cases to 6.9% (11 patients) thereafter (\( P = .12 \)). Complication rates showed improvement during the experience that did not reach statistical significance.

**Optimization of Oncologic Outcomes**

The number of lymph nodes harvested during resection to treat malignant neoplasms improved significantly after 80 cases from 17 to 26 (\( P < .001 \)) (Figure 2). The R1 resection rate showed improvement with experience that did not reach statistical significance. One subgroup of patients, cases 120 through 140, had an R1 resection rate of 26.3% that was more than twice the average rate for the remaining cohorts (7.4%; \( P = .02 \)). This group, however, tended to have larger tumors (3.3 cm vs 2.7 cm; \( P = .09 \)), and 2 of the 5 positive margins were read as R0 on intraoperative frozen section.

**Optimization of Resource Utilization Outcomes**

Outcomes related to efficacy and resource utilization included operative time, conversion to open surgery, length of stay, and readmission rate. The most significant improvement in operating times was seen after 80 procedures (mean SD), 581 [81] minutes for cases 1–80 vs 417 [78] minutes for cases 81–200; \( P < .001 \)) (Figure 3A). In the latter cohort, an incremental but statistically meaningful reduction in operating time was also observed between cases 81 through 140 and 141 through 200 (mean SD), 444 [73] minutes vs 389.6 [75] minutes; \( P < .001 \). To confirm this apparent improvement, a cumulative sum chart was constructed demonstrating a sharp decline in operative time after 80 cases, with an initial plateau phase followed by a third phase of steady improvement (Figure 3B). These 3 phases had distinct linear trends. In addition, the C chart in Figure 3A demonstrates that, after the initial 80 cases, variations in operative time were considerably less common, with only 3 patients having an operative time greater than 1 SD from the mean (SD, 113 minutes). The conversion rate to open surgery showed a steep decline after 20 procedures (35.0% vs 3.3%; \( P < .001 \)) (Figure 4). Reasons for conversion were failure to progress (\( n = 5 \)), unexpected vascular involvement (\( n = 5 \)), bleeding (\( n = 2 \)), and inability to tolerate pneumoperitoneum (\( n = 1 \)). Both the median length of stay and the number of readmissions decreased after 80 cases, but these outcomes did not reach statistical significance. The reasons for readmission are listed in the eTable in the Supplement and did not appear to change significantly throughout the experience.

**Ninety-Day Outcomes of 120 RPDs Following Optimization**

The quality analysis above demonstrated optimization of safety, efficiency, and oncologic capability in the first 80 cases. The outcomes of the subsequent 120 cases are shown in Table 2 and represent outcomes attained once proficiency and efficiency with the robotic technology had been achieved. When collectively comparing the preoperative demographics, diagnosis, and preoperative tumor sizes between the first 80 and last 120 cases, no differences were detected, suggesting that the improved outcomes observed in the latter cohort were not a result of selection bias.

**Discussion**

Despite significant improvement in recent decades, pancreatic resections continue to result in substantial morbidity.14 Surg-
Robotic innovation and the use of new minimally invasive technology may improve outcomes, particularly for PD.10 During implementation of RPD, safety outcomes must be rigorously scrutinized and the learning curve identified to safely disseminate this platform and allow comparisons with other techniques.
In this study, the learning curve needed to attain proficiency and efficiency for RPD was approximately 80 cases. With the exception of increased use of neoadjuvant therapy, patient demographics did not change significantly during our study period, suggesting that minimal selection bias was present to influence our results. This learning curve...
was representative of a pancreatic surgery team rather than a single attending surgeon. Optimization of outcomes required acquisition of skills during our experience by 4 attending surgeons (A.H.Z., A.J.M., M.E.H., and H.J.Z.), rotating residents and fellows, and a dedicated operative staff, including circulating nurses, surgical assistants, and anesthesiologists. All initial 80 cases required 2 attending surgeons for the duration of the procedure to ensure patient safety. After the initial 80 cases, a 2–attending surgeon approach was only required for the uncinate dissection and for cases deemed to be more complex or high risk by the surgeons. Given the complexity of using the robotic platform for PD, we believe this type of team approach was crucial to ensuring patient safety and outcome quality. While this method does not allow the identification of the individual learning curve for a single surgeon, we advocate implementing this type of team approach for surgeons who are adopting RPD. Trainees played an increasing role throughout the experience and were given increasing responsibility for key portions of the procedure based on their demonstrated proficiency, first with resection of the gallbladder, then with gastrojejunostomy (2-layered hand-sewn), followed by hepaticojejunostomy, and finally resection. Despite the participation of 4 primary attending surgeons throughout the experience, there was no difference in any of the outcomes based on the attending surgeon or the involvement of trainees. Interestingly, our learning curve findings closely resemble those previously reported for open PD by Tseng et al15 and Schmidt et al,16 suggesting that an intrinsic number of hours are needed to achieve familiarity with the anatomy and platform.

Our results indicate that estimated blood loss and conversion were the first metrics to demonstrate early optimization (approximately 20 cases). Not surprisingly, these 2 metrics are expected to improve simultaneously; when the surgeon learns to handle intraoperative bleeding successfully, conversion is often averted. In the case of the RPD, this outcome represents the use of fine suture (5-0 and 6-0 monofilaments) to control venous or arterial bleeding during the head and uncinate dissection. Another important metric, pancreatic fistula, also improves early (approximately 40 procedures) as the surgeon rapidly acquires the skills of suturing complex anastomoses.

Overall efficiency, as measured by operative time, took longer—up to 80 cases—owing to progressive refinement of the operative technique. The operative time beyond the learning curve (80 cases) was 417 minutes, which is comparable with open PD in several large studies.14,17 We postulate 4 reasons behind this long operative time learning curve. First, our surgical team implemented the robotic PD at an early stage of its inception and were thus “self-trained” as opposed to being mentored and guided on the optimal operative set-up and sequence. Second, we implemented numerous modifications to the operative technique during the first 60 to 80 cases to improve outcomes. Third, new surgeons were being integrated into the robotic program. Fourth, we were able to incorporate residents and fellows into the procedure and allow them to perform increasing portions of the operation. It is likely that fewer than 80 cases are required for the true optimization of operative time in the absence of some of these factors and in the presence of better guidance and mentorship for future surgeons who are adopting the procedure. The slope of our operative time (Figure 3B) suggests a continued gradual improvement, even after 80 procedures; it is possible that peak operative times for RPD may continue to improve with advancing technology, albeit at a slower rate compared with the learning curve. Lymph node harvesting outcomes also showed significant improvement after 80 cases. Although this outcome may be a function of better pathologic sampling, we presume this result to also be a reflection of the more technically efficient and comprehensive dissection of the adventitial tissue along the superior mesenteric artery that comes with increasing experience with the platform. Notably, other parameters, including complication rates and readmissions, improved beyond the learning curve but failed to reach statistical significance. This result was likely owing to an underpowered small sample size. These metrics are still within published outcomes of open PD.14,18

We established a process of quality outcome analysis during our implementation of RPD. Quantitative outcomes were plotted on control charts to allow easy visualization of outliers within a group as well as across the study period. Using this method, variations in outcomes were identified and root causes for deviation explored. All outcomes (beyond the learning curve of 80 cases) that deviated from the optimal performance metrics were scrutinized. Few outliers were detected with respect to outcomes beyond the learning curve, a reflection of the structured methods and reliable results with which the RPD can be performed once the learning curve is attained.

### Table 2. Optimized 90-Day Outcomes of the Last 120 Robotic Pancreatectoduodenectomies

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Value (n = 120)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operative time, mean (SD), min</td>
<td>417 (78)</td>
</tr>
<tr>
<td>Estimated blood loss, median (IQR), mL</td>
<td>250 (150-400)</td>
</tr>
<tr>
<td>Rate, %</td>
<td></td>
</tr>
<tr>
<td>Conversion</td>
<td>3.3</td>
</tr>
<tr>
<td>Transfusion</td>
<td>21.7</td>
</tr>
<tr>
<td>Mortality</td>
<td>3.3</td>
</tr>
<tr>
<td>Pancreatic fistula (ISGPF grade B/C)</td>
<td>6.9</td>
</tr>
<tr>
<td>Readmission</td>
<td>29.2</td>
</tr>
<tr>
<td>R0 resection</td>
<td>91.4</td>
</tr>
<tr>
<td>Clavien-Dindo classification rate, %</td>
<td></td>
</tr>
<tr>
<td>&lt;3</td>
<td>43.2</td>
</tr>
<tr>
<td>≥3</td>
<td>23.3</td>
</tr>
<tr>
<td>Length of stay, median (IQR), d</td>
<td>9 (7-14)</td>
</tr>
<tr>
<td>Lymph node harvest, median (IQR), No.</td>
<td>26 (19-32)</td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; ISGPF, International Study Group on Pancreatic Fistula.
Although we have shown that optimization of certain parameters, such as operative time, blood loss, and conversion to open surgery, can be achieved in a span of 80 cases, other important metrics, such as time to receipt of adjuvant therapy and survival, also need to be examined. These and other important outcomes can be evaluated and compared with open and laparoscopic platforms now that the learning curve for RPD has been identified. Cost is a central metric that also needs to be evaluated, and a comparison of the cost of RPD with open PD will be a major determinant in the adoption of this platform. This analysis can now be performed because the learning curve and optimized outcomes for RPD have been identified.

Conclusions

It appears that the learning curve for RPD is attained within 80 cases. In addition, our quality analysis of new surgical technology for complex operations, such as RPD, allows surgeon to monitor outcomes, ensuring patient safety, and provides valuable insight and guidelines for dissemination and adoption. This type of analysis could serve as a model for the implementation of new surgical procedures or technologies in the current outcomes-driven environment of quality health care.