Surgical Skill Assessment in Robot-Assisted Colorectal Cancer Resection using Operative Performance Indicators

Yiannis Haridemou

4th Year Project Report
Computer Science and Mathematics
School of Informatics
University of Edinburgh

2025

Abstract

This project investigates whether robot console log–derived Operative Performance Indicators (OPIs) can support automated assessment of technical performance in robot-assisted colorectal cancer surgery. Using MASTERY study data from 41 robot-assisted resections performed by 14 surgeons, task-level OPIs were aggregated per operation step and analyzed alongside postoperative outcomes including hospital length of stay and a composite patient-reported condition score derived from quality-of-life questionnaires using principal component analysis. The work characterizes typical task ordering, visualizes workflow variability across surgeons, identifies task–metric pairs strongly associated with outcomes, and demonstrates a surgeon-personalized feedback report that benchmarks key metrics using percentile ranges. Finally, supervised regression models were trained on selected OPIs to predict outcomes, illustrating the feasibility of forecasting recovery metrics from intraoperative behavior. The findings support OPIs as interpretable signals linked to patient recovery, while highlighting limitations related to cohort size, missing questionnaire responses, and skewed surgeon case distribution.

Research Ethics Approval

Ethics approval for the MASTERY study was granted by the Bloomsbury Research Ethics Committee, London (REC reference: 20/PR/0580, Date of REC opinion: 17 Dec 2020).

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Yiannis Haridemou*)

Acknowledgements

Special thanks to Professor Mohan Sridharan for the continued support and Professor Simon Bach for his valuable medical input.

Table of Contents

Chapter 1

Introduction

1.1 Motivation and Clinical Importance

Colorectal cancer is among the leading causes of cancer-related morbidity and mortality worldwide, representing a major public health concern with an increasing incidence in both developed and developing nations [1]. Surgical resection is the main curative treatment, especially in non-metastatic disease [2, 3]. The last two decades have witnessed a paradigm shift in surgical oncology with the adoption of minimally invasive techniques, particularly laparoscopic and robotic-assisted approaches, which aim to reduce postoperative complications and enhance recovery while maintaining oncological safety [4].

Robotic platforms such as the Da Vinci Surgical System have emerged as a widely used platform in this domain, offering enhanced dexterity, tremor filtration, and magnified three-dimensional visualization [5]. These features are particularly valuable in pelvic surgery, where access is constrained and dissection planes are narrow. In rectal cancer resections, robotic systems can facilitate precise mesorectal excision and nerve preservation, potentially improving both oncological and functional outcomes [6].

However, the benefits of robotic surgery are highly dependent on the skill of the operating surgeon. Numerous studies have demonstrated that surgeon experience and technical proficiency are critical determinants of outcomes such as operative time, blood loss, conversion to open surgery, margin status, complication rates, and even long-term recurrence [7, 8, 9].

1.2 Challenges in Assessing Surgical Skill

Assessing surgical skill remains a fundamentally difficult problem. Conventional tools like the Objective Structured Assessment of Technical Skills (OSATS) or Global Evaluative Assessment of Robotic Skills (GEARS) are reliant on human raters and are thus vulnerable to inconsistency, bias, and limited scalability. For instance, the evaluation may vary significantly depending on the seniority or familiarity of the assessor with the procedure or surgeon being evaluated [10].

The challenge is further compounded in robotic-assisted surgery, which introduces a complex workflow involving not only manual dexterity but also the strategic use of console features such as the camera, energy instruments, and clutch mechanisms [11].

Additionally, patient factors such as tumor location, BMI, prior surgeries, and disease severity introduce further variability. A surgeon operating on a difficult case may appear less efficient than another working on a straightforward one, even if their technical skill is comparable. This makes fair comparisons across surgeons or cases especially difficult without a robust framework that accounts for such heterogeneity [12].

1.3 Role of Robotic Surgery in Colorectal Procedures

Colorectal surgery encompasses a wide range of procedures including right hemicolectomy, left colectomy, low anterior resection, and abdominoperineal resection. In recent years, robotic systems have become particularly favored in complex resections such as total mesorectal excision (TME) for mid-to-low rectal cancers, where the confined pelvic space limits visibility and maneuverability using conventional laparoscopic tools [13, 14].

Robotic surgery offers several technical advantages over laparoscopy. The articulated wristed instruments allow for a greater range of motion than standard straight-stick tools, making dissection around critical structures like the hypogastric nerves more precise. The surgeon-controlled camera provides stable, high-definition 3D visualization, reducing reliance on assistants and improving depth perception during dissection [11].

The growing availability of system log data offers a unique opportunity to analyze not only what was done but how it was done. This study explores how these log-derived OPIs can reflect surgical technique and predict patient outcomes in robotic colorectal cancer resections.

1.4 Problem Statement

Despite the growing integration of robotic platforms in colorectal cancer surgery, a significant gap remains in how these systems are utilized for performance evaluation. Modern robotic systems like the da Vinci Xi automatically generate extensive log data for every procedure, including timestamps, instrument usage, camera movements, and energy applications. However, these data are rarely used beyond basic operational review or device troubleshooting [12].

This represents an underutilised opportunity to harness the precision and consistency of digital records to assess surgeon behavior in an objective and reproducible manner. Operative Performance Indicators (OPIs) provide a structured approach to quantifying technical skill by transforming low-level robotic console events—such as clutch activations, camera repositionings, and energy activations—into high-level behavioral metrics. By examining these metrics within specific surgical tasks and across a variety of cases, OPIs have the potential to reveal patterns of efficiency, consistency, and error that traditional assessments cannot capture [10, 12].

1.5 Study Goals

This study aimed to explore how operative performance indicators (OPIs) derived from robotic console data relate to patient recovery in colorectal cancer surgery. The research was divided into four phases, each targeting distinct objectives:

1. **Exploratory Workflow Analysis:** To characterize surgical task sequences and identify patterns across surgeons.
2. **Outcome Correlation Analysis:** To quantify relationships between specific OPIs and key postoperative outcomes, namely hospital length of stay and patient condition scores.
3. **Surgeon-Specific Evaluation:** To examine whether surgeon background influences outcomes and to generate individualized performance feedback reports.
4. **Predictive Modeling:** To assess the feasibility of forecasting outcomes using machine learning models trained on intraoperative metrics.

These phases collectively support the broader goal of data-driven skill assessment in robotic colorectal surgery.

Chapter 2

Background

2.1 Robotic Colorectal Surgery

The evolution of surgical techniques in colorectal cancer treatment has progressed from open surgery to laparoscopy and, more recently, to robotic-assisted interventions. While laparoscopic surgery reduced morbidity and recovery times compared to open procedures, its limitations—especially in complex pelvic dissections—prompted the development of robotic systems. The *da Vinci Xi Surgical System*, used in this study, improves upon laparoscopy by providing high-definition 3D visualization, wristed instruments with seven degrees of freedom, and ergonomic control interfaces that reduce surgeon fatigue. These attributes are especially beneficial in total mesorectal excision and other pelvic procedures, where precise dissection is critical and access is limited.

A unique feature of robotic platforms is their ability to generate detailed console event logs. These include timestamps for clutch usage, camera movements, energy activations, and kinematic data for each instrument. Unlike video-based studies that require manual annotation, these logs provide continuous, high-resolution input for objective surgical performance evaluation. Compared to prior studies, which often used aggregated summary statistics or limited sensor data, the present work leverages an unusually granular dataset with a high number of distinct robotic metrics across multiple colorectal surgical steps.

Robotic training follows a structured curriculum set by system manufacturers and surgical societies. Credentialing typically includes simulator modules, proctoring, and structured assessments. Nevertheless, skill variation persists, even among credentialed surgeons, emphasizing the need for automated, reproducible performance metrics that extend beyond subjective observation.

2.2 Operative Performance Indicators (OPIs)

Operative Performance Indicators (OPIs) are quantifiable features derived from robotic console logs that characterize intraoperative behavior. These metrics serve as proxies for skill, efficiency, and procedural fluency. Examples include clutch count (instrument repositioning), camera pan frequency (visual control), energy activation time (tissue dissection/cauterization), and derived kinematic quantities such as speed, acceleration, jerk, and wrist articulation patterns. These data-driven metrics allow for reproducible and scalable performance analysis across large datasets.

Several studies have validated the use of OPIs for evaluating surgeon skill. In their landmark work on robot-assisted prostatectomy, Hung et al. (2018) showed that machine learning models trained on OPIs could predict patient outcomes with moderate accuracy, such as blood loss and complication rates, while differentiating between experience levels based on automated metrics alone [15]. Hung et al. (2019) further investigated ”experts vs. super-experts” and demonstrated significant differences in OPIs between high-performing surgeons, suggesting that OPIs can capture nuances not easily discerned through traditional assessments [16].

A systematic review by Gerull et al. (2025) consolidated findings from multiple surgical specialties and concluded that OPIs are consistently associated with surgeon experience and technical performance, though most studies focused on simulation environments or small, homogeneous datasets [17]. Neis et al. (2024) applied OPIs to workflow segmentation in robot-assisted hysterectomy, highlighting the potential for task-specific analysis but relying on a narrower metric set and smaller sample sizes [18].

Relative to these studies, the present work diverges in two critical ways. First, it focuses on colorectal cancer resection—a domain less represented in prior OPI literature. Second, it introduces a richer set of robotic metrics at the task level, enabling fine-grained analysis across the distinct phases of colorectal surgery. Most previous studies operate at a case-level resolution; by contrast, this thesis analyzes correlations and predictive power of OPIs within discrete surgical steps, thus enhancing interpretability and feedback precision.

2.3 Related Research and Gaps

Traditional skill assessment frameworks such as OSATS, GOALS, and MISTELS rely on rater-based scoring systems. While these tools capture dimensions like bimanual dexterity, instrument handling, and efficiency, they are limited by inter-rater variability and the logistical burden of expert observation. OPIs, in contrast, enable automated, scalable evaluation with reduced subjectivity and higher temporal resolution.

The validation of OPIs across specialties has opened new avenues for skill analysis and predictive modeling. However, the literature remains sparse in several respects. First, colorectal surgery has received limited attention in OPI research, despite being one of the most technically challenging domains. Second, few studies align intraoperative metrics with postoperative outcomes. While Hung et al. (2018) and Neis et al. (2024) explore predictive modeling, they do not examine the full surgical workflow or patient-centered outcomes such as recovery time or self-reported quality of life.

Moreover, most prior studies utilize relatively coarse-grained data, often restricted to a handful of metrics such as clutch count or energy time. In contrast, this work integrates over 50 distinct metrics, including percentile-based and instrument-specific features, enabling a more nuanced characterization of surgeon behavior. It also incorporates both clinical outcomes (length of stay, complication forms) and validated patient-reported metrics (composite quality-of-life score from PCA), thereby extending the analytic scope beyond technical performance.

Finally, this thesis proposes and demonstrates a novel surgeon-specific feedback mechanism, offering percentile-based benchmarking grounded in outcomes-linked OPIs. By integrating technical metrics with outcome prediction and individualized reporting, this work contributes both methodological innovation and practical tools for skill development in robotic colorectal surgery.

Chapter 3

Data

3.1 Data Collection Overview

Data were collected as part of the MASTERY research study, conducted by the Royal College of Surgeons Robotics Working Group. The goals of this study are: (1) to improve surgeon training by defining actionable, task-based metrics that can benchmark technical performance and (2) to improve patient outcomes by using task-based metrics to determine optimal procedure workflow and surgeon technical efficiencies. [19] The following sections will outline the 3 types of datasets that have been used.

3.2 Data Cleaning

Despite each dataset having specialised filtering appropriate for its use, all datasets have been cleaned using universal filters. All rows that had all their metrics or qualitative variable columns empty, have been dropped. All columns that have been empty, or consisted of the same single repeated value have also been dropped. The reason for that being that no insight can be derived by columns with identical data, due to lack of variation.

3.3 Intraoperative Data

3.3.1 Qualitative Surgery Description and Evaluation

This data originates from post-operation forms completed by the surgeons. It includes overall and task-specific success or complications. It also indicates which of the 15 standardised tasks (including ”Other”) have been performed, and by whom (lead surgeon, second surgeon or trainee). The data has been filtered to include only the 41 relevant operations, that exist in the metrics dataset, described below.

3.3.2 Surgery Task-Specific Metrics

The data has been collected through the Intuitive Data Recorder (IDR) device during live robotic assisted surgery using the Intuitive Da Vinci Xi robot. [20] For each patient's operation, there is a row for each task performed in chronological order. For each task there exists start and end time, that has been annotated by a human, as well as the associated metrics values automatically produced by the IDR. The lead surgeon of a patient is also specified and there are 41 operations available. Since each task may be performed in a non-continuous manner, either interrupted by a break or another task, for each operations metrics should be aggregated so that there is one instance of each task. For instance, if total mesorectal excision happened in several non-continuous parts, then some metrics, such as total master clutch use, have been summed and some metrics, such as average duration across all individual energy events, have been averaged.

Below, a full list of the available metrics is included, along with their description and respective measurement unit.

`Variable`	**Description**
`system_dur_sec`	Time span between the start of the task to the end of the task (Seconds)
`evtcnt_master_clutch`	Number of times of master clutching during the surgical task, either through finger clutching or pedal clutching
`evtcnt_master_clutch_finger`	Number of times of master clutching just through finger clutching
`evtcnt_master_clutch_pedal`	Number of times of master clutching just through pedal clutching
`evtcnt_camera_control`	Number of times the camera was moved by pressing the pedal
`evtcnt_head_out`	Number of times the surgeon took their head out of the console
`evtcnt_arm_swap`	Number of times control was swapped to a different robotic arm
`evtcnt_energy_activation`	Number of instrument energy activations using pedals
`energy_avgduration`	Average duration per energy event (Seconds)
`energy_totalduration`	Total time energy was applied (Seconds)
`eom_usm#`	Total linear distance traveled by the tip of USM# in any direction (Meters)
`activetime_any`	Total time any instrument was moving (not idle) (Seconds)
`activetime_usm#`	Time the instrument on USM# was moving (not idle) (Seconds)
`speed_pct_usm#_X%`	Instrument tip speed at the Xth percentile while moving (Meters per second)
`accel_pct_usm#_X%`	Instrument tip acceleration at the Xth percentile (Meters per second^2)

Variable	Description
jerk_pct_usm#_X%	Instrument tip jerk at the Xth percentile (Meters per second^3)
wrist_usm#_roll	Total angular movement by the roll joint of USM#'s wrist (Radians)
wrist_usm#_pitch	Total angular movement by the pitch joint of USM#'s wrist (Radians)
wrist_usm#_yaw	Total angular movement by the yaw joint of USM#'s wrist (Radians)
wrist_activetime_usm#	Time the wrist of USM# was moving (Seconds)
wrist_activetime_any	Time any wrist joint (roll, pitch, yaw) was moving for any USM (Seconds)
eom_ssc	Total distance moved by hand controllers across all consoles (Meters)
ssc_eom_mtmY#	Total distance moved by hand Y (left/right) on console # (Meters)
wristmove_ssc	Total wrist movement across all consoles (Radians)
ssc_activetime_mtmY#	Time hand Y was actively moving on console # (Seconds)
wrist_mtmY#_yaw	Yaw joint rotation of the hand controller on console # (Radians)
wrist_mtmY#_roll	Roll joint rotation of the hand controller on console # (Radians)
wrist_mtmY#_pitch	Pitch joint rotation of the hand controller on console # (Radians)
speed_pct_yaw_mtmY#_X%	Yaw angular speed at Xth percentile for hand controller Y on console # (Radians per second)
speed_pct_pitch_mtmY#_X%	Pitch angular speed at Xth percentile for hand controller Y on console # (Radians per second)
speed_pct_roll_mtmY#_X%	Roll angular speed at Xth percentile for hand controller Y on console # (Radians per second)
wrist_speed_pct_yaw_usm1_1	Yaw angular speed of USM1 wrist at 1st percentile (Radians per second)
wrist_speed_pct_roll_usm1_1	Roll angular speed of USM1 wrist at 1st percentile (Radians per second)
wrist_speed_pct_pitch_usm1_1	Pitch angular speed of USM1 wrist at 1st percentile (Radians per second)

3.4 Surgeon Data

There is a Surgeon Registration Form that includes data, such as the years of experience of the surgeon (general, not particular for robotic surgery), their annual and total case volume (only robot-assisted) and completion or current attendance of a robotic fellowship. A robotic fellowship is a clinical experience for students to gain robotic-assisted surgical skills, after completing a conventional medical residency program.

[21] There are total 14 surgeons corresponding to the available 41 operations. A key limitation is that the distribution of cases across surgeons is highly imbalanced: one surgeon ("70-4-JSK") performed 22 of the 41 operations, while the remaining 13 surgeons performed the other 19 operations.

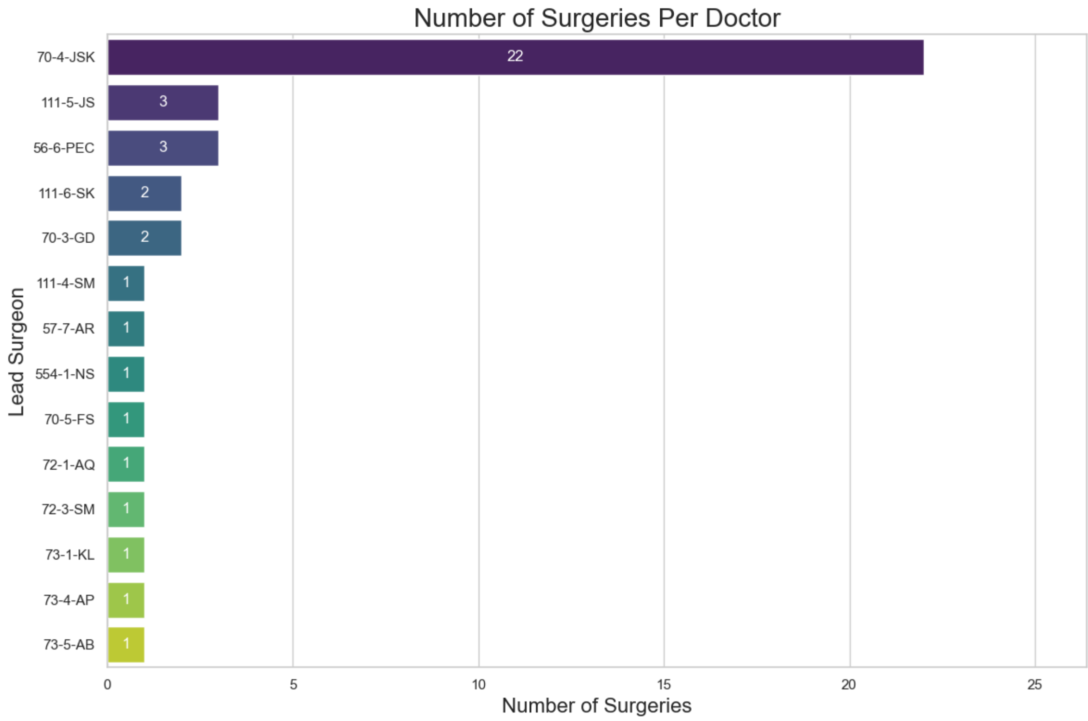

Figure 3.1: Distribution of operations per surgeon in the cohort (showing dataset imbalance).

3.5 Patient Data

3.5.1 Patient Characteristics

The patient cohort included adults with histologically confirmed colorectal cancer who underwent robot-assisted resection between 2020 and 2023. Collected variables included age, BMI, comorbidities and relevant risk factors. Cases involving emergency conversions to open surgery or palliative resections were excluded, as the outcomes would not be directly related to the metrics of the surgical robot.

3.5.2 30 Days Post Operation Form

Postoperative outcomes were captured using standardized forms at 30 days, documenting complications, readmission status, and discharge disposition and length of stay. Outcomes were coded as binary, continuous or discrete variables, depending on the clinical field.

3.5.3 Patient Quality of Life Questionnaire

Quality of life was self-reported using a colorectal-specific questionnaire administered postoperatively. It includes only questions with more than 30 responses out of 37 to ensure variable reliability. The questionnaire contains variables about pain, fatigue, mobility, appetite, and other physical symptoms. Responses were later aggregated via principal component analysis to yield a composite patient condition score.

Chapter 4

Methodology and Results by Research Phase

4.1 Phase 1: Exploratory Data Analysis

4.1.1 Average Task Sequence

The average task sequence of colorectal resection has been determined, by calculating a position score for every task. This approach has been used, since each surgery has a different amount of tasks, repetition or lack of certain tasks. This sequence has been used in subsequent graphs.

Seq.	Task	Avg. Score (0–100)	Surgeries	Freq.
1	Initial Exposure	13.80	28	0.68
2	Medial To Lateral Mobilization Of The Rectum, Sigmoid, And Descending Colon	18.53	40	0.98
3	Splenic Flexure Mobilization	46.24	33	0.80
4	Ligation And Division Of The Named Pedicle	46.70	39	0.95
5	Total Mesorectal Excision	48.46	39	0.95
6	Lateral To Medial Mobilization Of Sigmoid And Descending Colon	51.71	40	0.98
7	Hemostasis	60.02	5	0.12
8	Transection Of The Rectum (Distal End Of Specimen)	66.65	30	0.73
9	Preparation Of The Distal End Of The Specimen For Transection	67.98	22	0.54
10	Preparation Of The Proximal End Of The Specimen For Transection	68.66	23	0.56
11	Colorectal Anastomosis	82.95	27	0.66
12	Transection Of The Descending Colon (Proximal End Of Specimen)	90.56	5	0.12

4.1.2 Step Transition Probability Heatmap

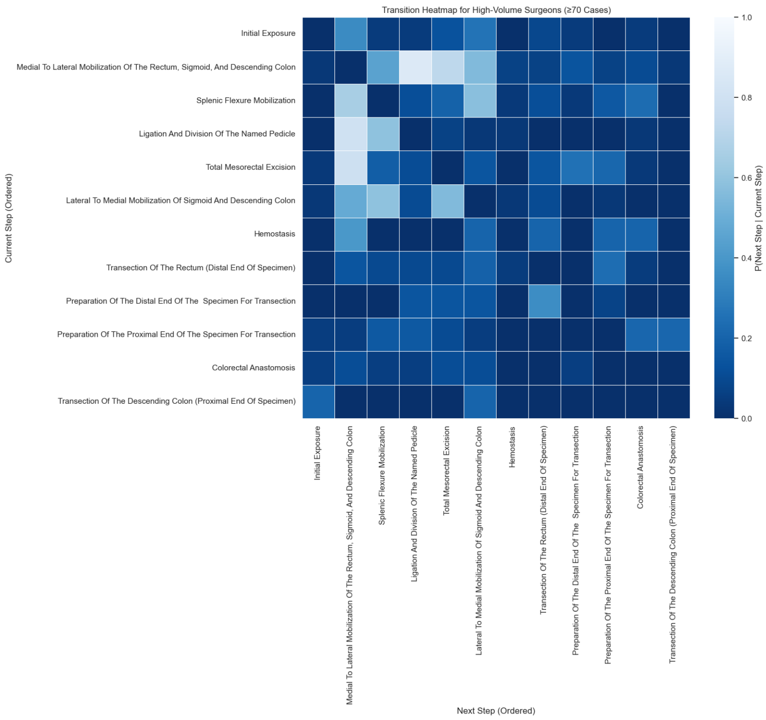

Figure 4.1: Step-transition probability heatmap for the high case-volume surgeon group (grouping as defined in this study).

Figure 4.2: Step-transition probability heatmap for the low case-volume surgeon group (grouping as defined in this study).

4.1.3 Surgical Workflows & Gantt Analysis

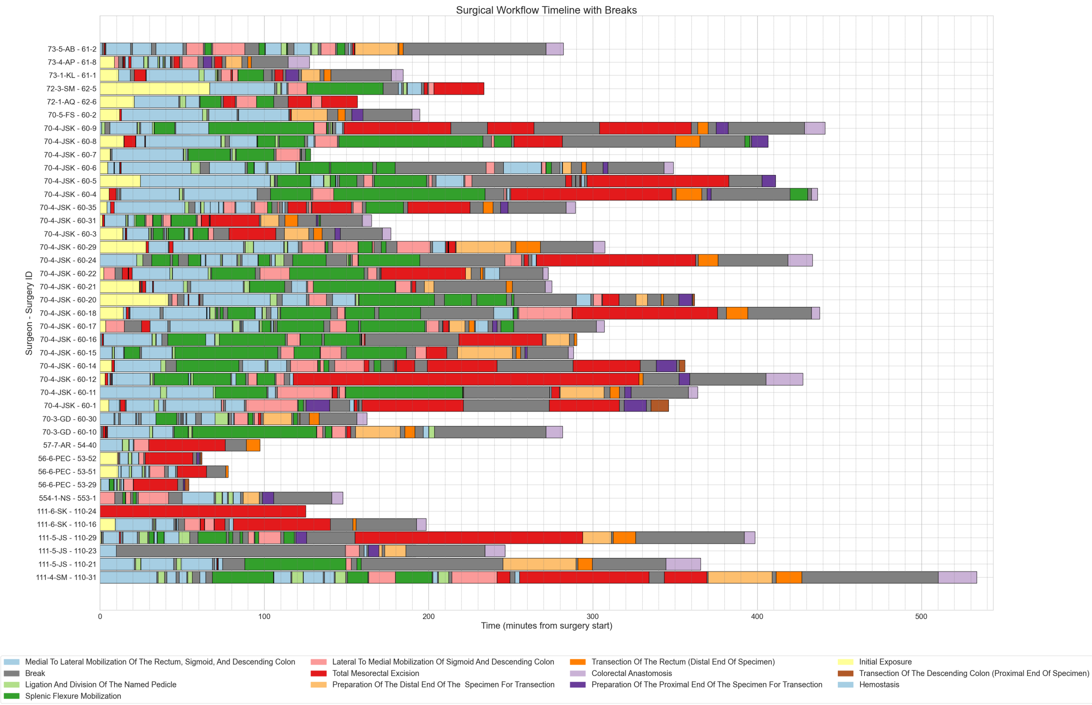

Figure 4.3: Gantt chart of task execution across operations, grouped by surgeon (each row corresponds to one operation).

From the step transition probability heatmaps for high and low case volume surgeons, as well as the detailed Gantt chart showing the workflow of each operation, grouped by surgeon, it can be seen that each surgeon takes a vastly different approach. However, in later phases no clear relationship was observed between case volume and the two outcome measures used in this study.

4.2 Phase 2: Relationship of Metrics to Outcomes

4.2.1 Hospital Length of Stay

4.2.1.1 Methods

This phase aimed to identify task-specific robotic metrics significantly correlated with postoperative hospital length of stay.

The analysis integrated operative performance indicators (OPIs) with clinical outcome data by merging aggregated da Vinci Xi robotic console metrics with 30-day postoperative patient records. Metrics considered included various temporal, kinematic, and instrument usage features, such as clutch events, wrist movements, and tool speed distributions. Metrics corresponding to the first quartile (P25) and third quartile (P75) were excluded to maintain focus on central and extreme value behaviors (e.g., minimum, median, maximum).

Each metric was analyzed separately for individual surgical steps. For each task, Pearson correlation coefficients were computed between OPIs and total hospital stay. Only task-metric pairs with at least three complete data points were considered, and results with strong correlation ($|r| \geq 0.5$) and statistical significance ($p \leq 0.05$) were retained for interpretation.

4.2.1.2 Results

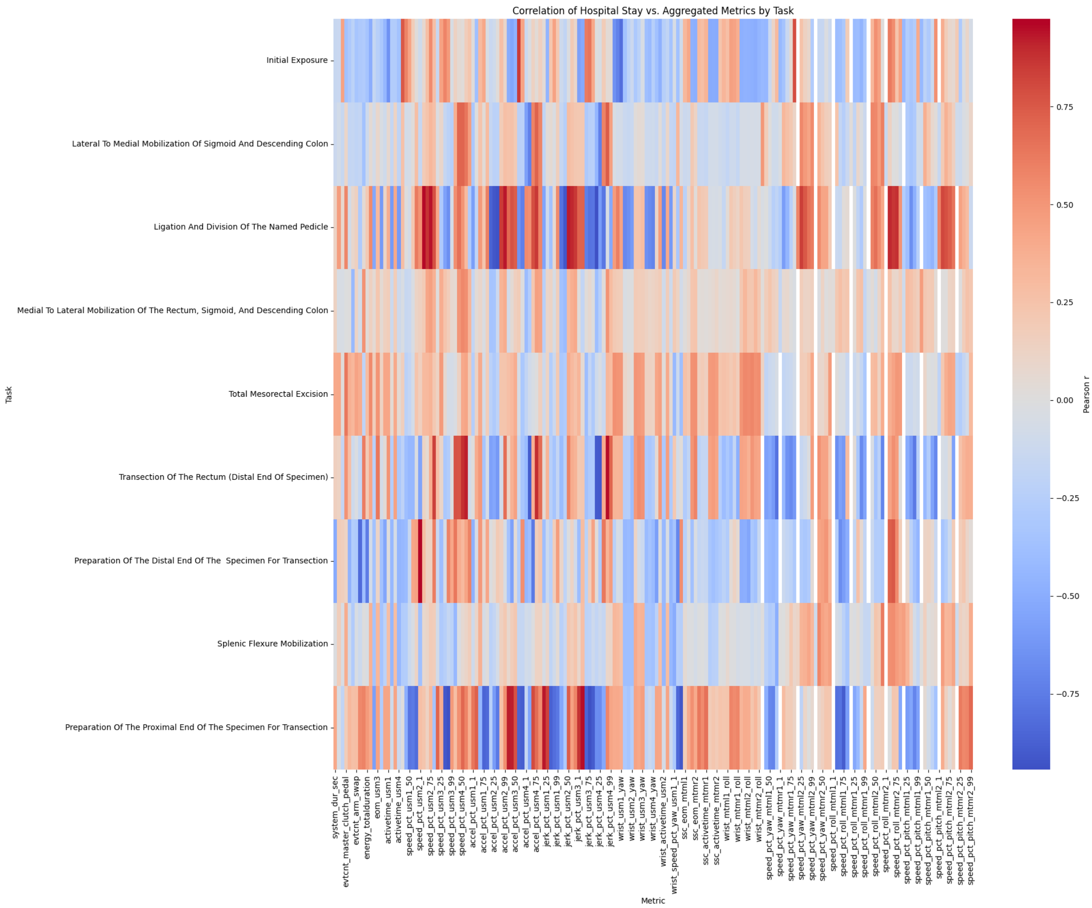

Figure 4.4: Task-level Pearson correlations between OPIs and hospital length of stay (LOS) for retained metric–task pairs under the study's filtering criteria.

The analysis identified a substantial number of OPIs that were strongly and significantly correlated with hospital length of stay, varying across surgical tasks. Notably, the task *Ligation and Division of the Named Pedicle* showed the highest density of significant metrics. For instance, `accel_pct_usm2_50` exhibited a near-perfect negative correlation ($r = -0.93$, $p < 0.001$), suggesting that smoother acceleration profiles in this step were associated with shorter hospital stays. Similarly, high correlations were observed with jerk-based metrics, such as `jerk_pct_usm2_50` ($r = 0.92$, $p < 0.01$).

Strong correlations were also found in tasks like *Initial Exposure*, *Total Mesorectal Excision*, and *Transection of the Rectum*, involving metrics related to wrist articulation and energy application duration. For example, increased average energy activation duration during rectal mobilization correlated positively with longer stays ($r = 0.52$, $p = 0.013$).

Overall, the findings indicate that more efficient and controlled tool handling during key dissection and ligation steps is linked to improved short-term recovery, as reflected in reduced hospital length of stay.

Metric	**Task**	**r**	**p**
speed_pct_usm1_1	Initial Exposure	0.751	0.0198
wrist_usm1_pitch	Initial Exposure	-0.719	0.0290
wrist_usm1_yaw	Initial Exposure	-0.814	0.0075
speed_pct_yaw_mtmr1_99	Initial Exposure	0.792	0.0109
speed_pct_usm4_50	Lateral To Medial Mobilization	0.709	0.0014
accel_pct_usm4_1	Lateral To Medial Mobilization	-0.574	0.0160
accel_pct_usm4_50	Lateral To Medial Mobilization	0.587	0.0133
accel_pct_usm4_99	Lateral To Medial Mobilization	0.585	0.0137
jerk_pct_usm4_50	Lateral To Medial Mobilization	0.585	0.0137
speed_pct_yaw_mtml1_1	Lateral To Medial Mobilization	0.501	0.0407
speed_pct_roll_mtml2_99	Lateral To Medial Mobilization	0.604	0.0103
accel_pct_usm2_50	Ligation Of The Named Pedicle	-0.934	0.0007
jerk_pct_usm4_1	Ligation Of The Named Pedicle	-0.943	0.0004
speed_pct_usm2_50	Ligation Of The Named Pedicle	0.870	0.0049
jerk_pct_usm2_50	Ligation Of The Named Pedicle	0.922	0.0011
wrist_usm4_yaw	Ligation Of The Named Pedicle	-0.714	0.0466
speed_pct_pitch_mtml2_50	Ligation Of The Named Pedicle	0.782	0.0218
energy_avgduration	Medial To Lateral Mobilization	0.520	0.0131
evtcnt_master_clutch_pedal	Total Mesorectal Excision	0.632	0.0037
wrist_usm1_roll	Total Mesorectal Excision	0.530	0.0197
wrist_mtmr2_pitch	Total Mesorectal Excision	0.531	0.0193
speed_pct_usm2_99	Transection Of The Rectum	0.833	0.0102
speed_pct_usm4_50	Transection Of The Rectum	0.877	0.0043

Metric	Task	r	p
jerk_pct_usm4_1	Transection Of The Rectum	-0.906	0.0020
evtcnt_energy_activation	Preparation Of The Distal End	-0.827	0.0218
speed_pct_usm2_1	Preparation Of The Distal End	0.974	0.0002
accel_pct_usm4_50	Preparation Of The Distal End	-0.757	0.0487
wrist_speed_pct_yaw_usm1_1	Splenic Flexure Mobilization	-0.504	0.0199
speed_pct_roll_mtml2_99	Splenic Flexure Mobilization	0.615	0.0331
speed_pct_usm3_50	Preparation Of The Proximal End	-0.911	0.0115
jerk_pct_usm1_50	Preparation Of The Proximal End	-0.872	0.0234
wrist_speed_pct_pitch_usm1_	Preparation Of The Proximal End	-0.918	0.0099

4.2.2 Patient Condition Score

4.2.2.1 Methods

This phase aimed to construct and validate a composite measure of post-operative patient well-being using quality-of-life questionnaire responses.

To reduce dimensionality and derive an interpretable condition score, principal component analysis (PCA) was applied to 49 patient-reported outcome variables collected postoperatively. These variables were drawn from a colorectal-specific quality-of-life questionnaire covering functional limitations, pain, psychological symptoms, bowel function, and body image. All responses were standardized to zero mean and unit variance prior to PCA.

The first principal component (PC1), which explained approximately 32% of the total variance, was selected to represent the dominant axis of variation in post-operative condition. Questionnaire items with high positive loadings on PC1 included indicators of fatigue, interference with daily activities, emotional distress, and physical discomfort. A condition score for each patient was computed as the projection of their standardized questionnaire responses onto PC1.

Internal consistency of the full item set was assessed using Cronbach's alpha, yielding a high reliability coefficient ($\alpha = 0.94$), confirming the coherence of the questionnaire as a unidimensional construct.

To investigate which robotic metrics during specific surgical tasks correlated with patient recovery, Pearson correlations were computed between operative metrics and the PCA-based condition score, stratified by task. Only strong ($|r| \geq 0.5$) and statistically significant ($p \leq 0.05$) associations were reported.

4.2.2.2 Results

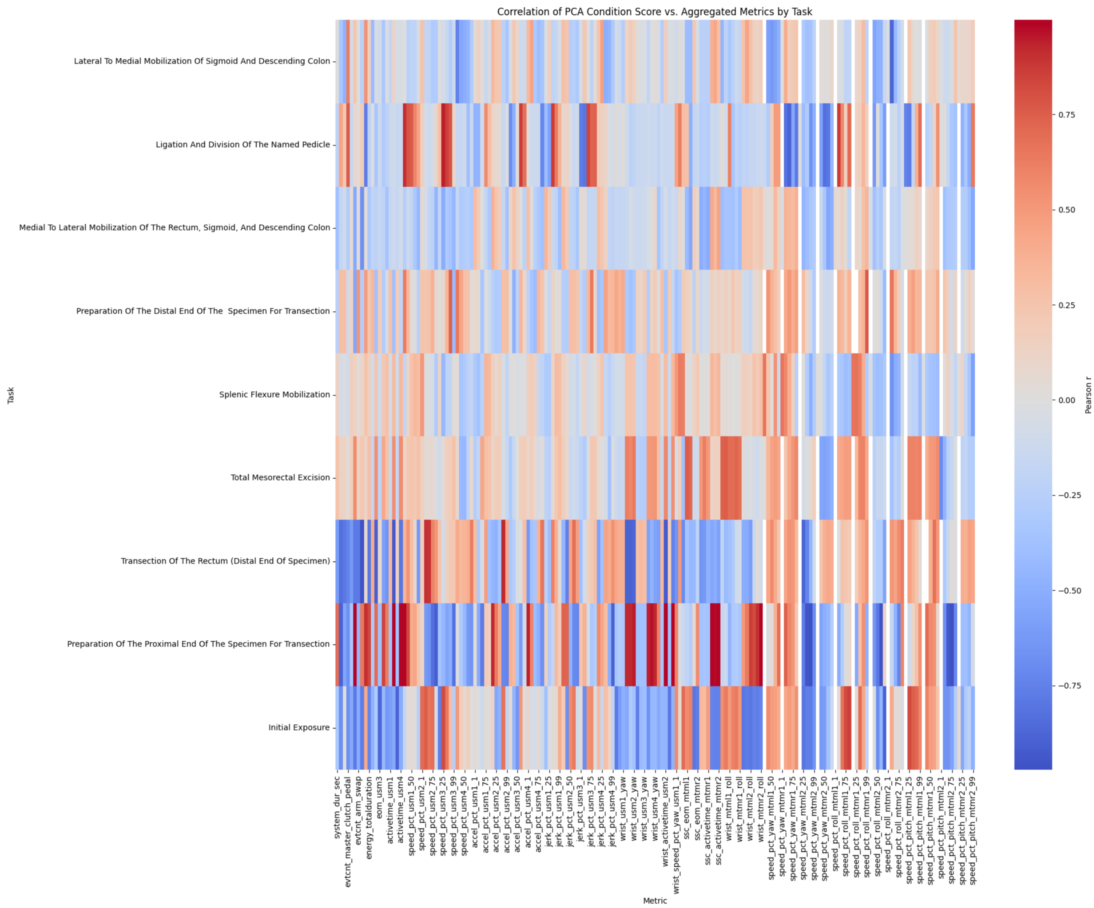

Figure 4.5: Task-level Pearson correlations between OPIs and the PCA-derived patient condition score (PC1) for retained metric–task pairs under the study's filtering criteria.

Several strong correlations were observed between task-specific operative metrics and patient condition scores. Notably, during *Transection of the Rectum*, multiple metrics demonstrated near-perfect correlations. For instance, `activetime_usm2` was negatively associated with condition score ($r = -0.94$, $p < 0.05$), suggesting that longer tool activation during transection corresponded to worse patient-reported outcomes. Similarly, wrist movement features such as `wrist_usm2_pitch` and `wrist_usm2_roll` exhibited strong negative correlations, indicating potential overmanipulation as a marker of suboptimal technique.

In contrast, metrics during the *Total Mesorectal Excision* phase showed positive correlations, such as `ssc_eom_mtml1` ($r = 0.76$, $p < 0.05$), implying that more efficient or balanced tool usage was linked to better patient recovery.

Overall, the results suggest that certain robotic performance metrics may serve as indicators of patient-centered outcomes, particularly in technically demanding phases like transection and mesorectal dissection.

Metric	**Task**	**r**	**p**
speed_pct_usm4_1	Lateral To Medial Mobilization	-0.749	0.0202
evtcnt_energy_activation	Medial To Lateral Mobilization	-0.676	0.0158
wrist_speed_pct_pitch_usm1_1	Splenic Flexure Mobilization	0.624	0.0301
speed_pct_yaw_mtmr1_1	Splenic Flexure Mobilization	0.661	0.0191
speed_pct_pitch_mtmr2_99	Splenic Flexure Mobilization	-0.583	0.0467
ssc_eom_mtml1	Total Mesorectal Excision	0.756	0.0114
wrist_mtmr1_pitch	Total Mesorectal Excision	0.723	0.0181
speed_pct_pitch_mtml2_1	Total Mesorectal Excision	-0.699	0.0245
evtcnt_energy_activation	Transection Of The Rectum	-0.971	0.0060
eom_usm2	Transection Of The Rectum	-0.908	0.0329
activetime_usm2	Transection Of The Rectum	-0.940	0.0173
speed_pct_usm2_50	Transection Of The Rectum	0.893	0.0412
wrist_usm2_pitch	Transection Of The Rectum	-0.887	0.0449
evtcnt_head_out	Preparation Of The Proximal End	0.996	0.0036
activetime_usm2	Preparation Of The Proximal End	0.990	0.0096
speed_pct_usm1_1	Preparation Of The Proximal End	0.998	0.0020
wrist_activetime_usm4	Preparation Of The Proximal End	0.985	0.0155
ssc_activetime_mtmr2	Preparation Of The Proximal End	0.990	0.0101
wrist_mtmr2_roll	Preparation Of The Proximal End	0.989	0.0105
activetime_usm3	Initial Exposure	-0.885	0.0463
speed_pct_roll_mtml2_99	Initial Exposure	-0.897	0.0392

4.3 Phase 3: Surgeon-Specific Analysis

4.3.1 Relationship of Surgeon Characteristics to Outcomes

4.3.1.1 Methods

To evaluate whether standard indicators of surgeon experience are predictive of patient outcomes, a correlation analysis was conducted between surgeon characteristics and postoperative metrics. The independent variables included fellowship completion status, years in practice, total number of indexed robotic colorectal cases, and annual case volume. These were correlated against two outcome measures: hospital length of stay and a composite patient condition score derived from principal component analysis. Pearson correlation coefficients were used to assess linear associations, and only results from complete cases were retained for each comparison.

4.3.1.2 Results

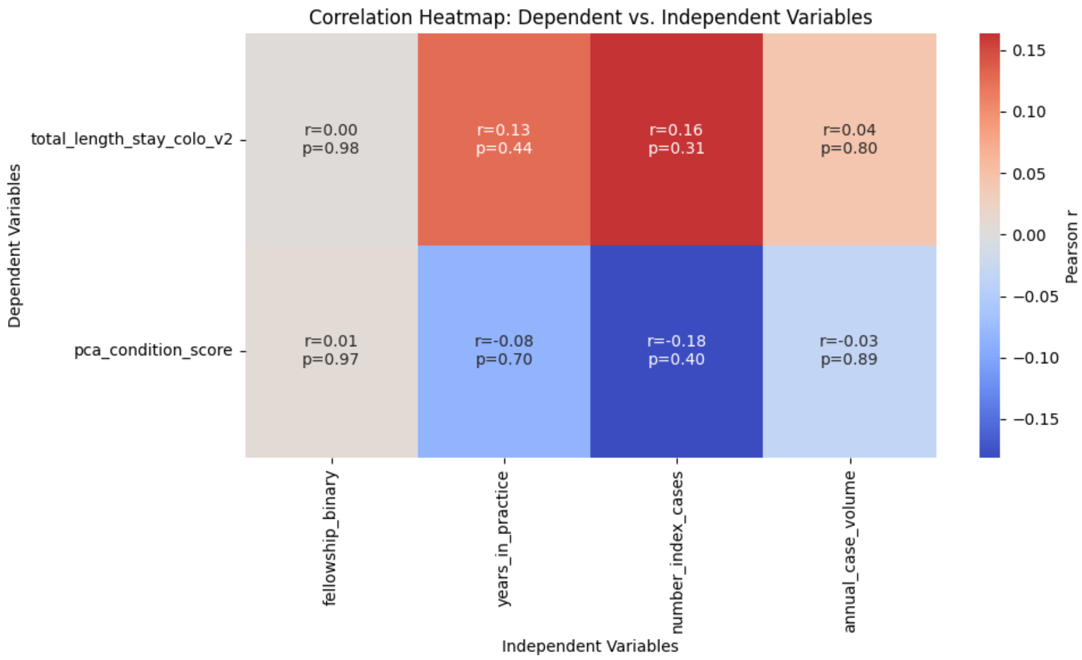

Figure 4.6: Pearson correlations between surgeon-level characteristics and postoperative outcomes (LOS and condition score) in this cohort.

None of the tested surgeon-level characteristics exhibited statistically significant correlations with either length of hospital stay or postoperative condition scores. Correlation coefficients were uniformly weak (e.g., $r = 0.13$ for years in practice vs. length of stay, $r = -0.03$ for annual case volume vs. condition score), with all p-values exceeding 0.3. These findings suggest that conventional surgeon experience metrics may not adequately capture the variation in outcomes observed in this cohort, highlighting the potential value of performance-based measures such as OPIs for more granular skill assessment.

4.3.2 Surgeon Personalised Feedback Report Generation

4.3.2.1 Methods

A feedback system was developed to generate individualized reports for surgeons based on their operative performance indicator (OPI) metrics. This system leverages previously identified metric-task pairs that exhibited significant correlations with clinical outcomes—namely, hospital length of stay and postoperative condition scores. For each of these metrics, percentile distributions were computed from the aggregated dataset to create population-level benchmarks.

To generate a feedback report, the system compares a surgeon's performance on each key metric to the benchmark distribution, identifying where their value falls in the percentile range (e.g., between the 40th and 50th percentile). This comparison is then contextualized with the directionality of the metric's correlation: for example, a surgeon with above-average energy activation duration in a task where longer activation correlates with poorer outcomes would receive constructive feedback. The entire process is automated to produce a human-readable PDF summarizing key performance insights.

4.3.2.2 Results

The generated report provides percentile-based feedback for each critical metric-task combination relevant to clinical outcomes. Surgeons can expect messages indicating where their performance stands relative to peers, along with interpretations of how those values relate to patient recovery or physical condition. This format supports data-driven reflection and skill development.

In the following page the first page out of ten of an example feedback report can be viewed.

Surgeon Feedback Report

Patient ID: 60-17
Lead Surgeon: 70-4-JSK

1. In task *Ligation And Division Of The Named Pedicle*, your `accel_pct_usm2_50` is between the 80th and 90th percentile. `accel_pct_usm2_50` is positively correlated with Patient Length of Stay (higher is worse).

2. In task *Ligation And Division Of The Named Pedicle*, your `accel_pct_usm3_50` is between the 20th and 30th percentile. `accel_pct_usm3_50` is positively correlated with Patient Length of Stay (higher is worse).

3. In task *Lateral To Medial Mobilization Of Sigmoid And Descending Colon*, your `accel_pct_usm4_50` is between the 50th and 60th percentile. `accel_pct_usm4_50` is positively correlated with Patient Length of Stay (higher is worse).

4. In task *Ligation And Division Of The Named Pedicle*, your `accel_pct_usm4_50` is between the 40th and 50th percentile. `accel_pct_usm4_50` is positively correlated with Patient Length of Stay (higher is worse).

5. In task *Preparation Of The Distal End Of The Specimen For Transection*, your `accel_pct_usm4_50` is between the 80th and 90th percentile. `accel_pct_usm4_50` is negatively correlated with Patient Length of Stay (higher is better).

6. In task *Preparation Of The Proximal End Of The Specimen For Transection*, your `activetime_usm2` is between the 70th and 80th percentile. `activetime_usm2` is positively correlated with Patient Physical Condition Score (higher is worse).

7. In task *Transection Of The Rectum (Distal End Of Specimen)*, your `activetime_usm2` is between the 60th and 70th percentile. `activetime_usm2` is negatively correlated with Patient Physical Condition Score (higher is better).

4.4 Machine Learning for Outcome Prediction

4.4.1 Hospital Length of Stay Prediction

4.4.1.1 Methods

To evaluate the predictive utility of operative performance indicators (OPIs) for estimating hospital length of stay (LOS), a supervised machine learning approach was employed. A curated set of OPI features was selected based on statistically significant correlations with LOS identified in earlier phases. These features spanned multiple tasks and sensor-derived motion characteristics (e.g., speed, acceleration, jerk, and instrument angles).

The dataset was cleaned to include only complete patient records across all selected metrics and LOS labels. After preprocessing, the data was randomly split into training and testing sets in an 80/20 ratio. Four regression models were developed and evaluated: linear regression (baseline), random forest, support vector regression (SVR), and XGBoost. Each model was integrated into a pipeline with standardized feature scaling and hyperparameter tuning via grid search and five-fold cross-validation. Model performance was assessed using standard regression metrics: mean absolute error (MAE), root mean squared error (RMSE), and R^2 score.

4.4.1.2 Results

All models were evaluated on a test set of 24 samples. The baseline linear regression model yielded an MAE of 4.20 days and RMSE of 5.61. In contrast, non-linear models demonstrated improved accuracy. The random forest regressor achieved an MAE of 2.92 and RMSE of 4.13, while SVR recorded slightly better MAE (2.77) but slightly worse RMSE (4.20). The best overall performance was observed with XGBoost, which achieved an MAE of 2.85 and RMSE of 3.97. These results suggest that LOS can be predicted with moderate accuracy from selected OPIs, supporting their relevance not only for post-hoc assessment but also for potential real-time outcome estimation. The consistent improvement of tree-based models over linear baselines highlights the non-linear nature of the relationship between robotic metrics and clinical outcomes.

Metric	Value
Linear Regression MAE	4.198
Linear Regression MSE	31.478
Linear Regression RMSE	5.611
Random Forest MAE	2.923
Random Forest MSE	17.021
Random Forest RMSE	4.126
Support Vector Regressor MAE	2.772
Support Vector Regressor MSE	17.623
Support Vector Regressor RMSE	4.198
XGBoost MAE	2.851
XGBoost MSE	15.738
XGBoost RMSE	3.967
Test Samples	24
Train Samples	96
Total Samples	120
Test Size	0.20
Test Set Std. Deviation	3.869

4.4.2 Patient Condition Score Prediction

4.4.2.1 Methods

The methodology for predicting patient condition scores followed the same supervised machine learning pipeline used for hospital length of stay, with two key differences.

First, the target variable was a principal component analysis (PCA)-based composite condition score derived from postoperative quality-of-life questionnaires. Second, the features included in model training were metrics previously identified as significantly correlated with this PCA score. Data was filtered to retain complete entries across both predictor metrics and target values, resulting in a smaller dataset. All four regression models—linear regression, random forest, support vector regression (SVR), and XGBoost—were trained with the same hyperparameter tuning strategy and evaluation metrics as the LOS prediction task.

4.4.2.2 Results

The linear regression baseline yielded the highest prediction error, with a mean absolute error (MAE) of 5.84 and root mean squared error (RMSE) of 7.65. The SVR model demonstrated the best performance, achieving an MAE of 1.96 and RMSE of 3.37. XGBoost and random forest followed with RMSE values of 3.86 and 4.16, respectively. As the standard deviation of the target variable in the test set was 3.39, the SVR model's RMSE indicates a moderate level of predictive accuracy. While these results suggest that selected operative metrics capture meaningful patterns related to postoperative condition, the PCA-derived nature of the target score—being unitless and compositional—introduces interpretive limitations. Performance metrics should thus be contextualized relative to the variance of the score, rather than compared directly to those from other outcome predictions such as length of stay.

Metric	Value
Linear Regression MAE	5.839
Linear Regression MSE	58.472
Linear Regression RMSE	7.647
Random Forest MAE	2.925
Random Forest MSE	17.271
Random Forest RMSE	4.156
Support Vector Regressor MAE	1.963
Support Vector Regressor MSE	11.354
Support Vector Regressor RMSE	3.370
XGBoost MAE	2.561
XGBoost MSE	14.895
XGBoost RMSE	3.859
Test Samples	14
Train Samples	52
Total Samples	66
Test Size	0.20
Test Set Std. Deviation	3.393

Chapter 5

Discussion

5.1 Summary of Findings

This study demonstrates that intraoperative robotic metrics—referred to as Operative Performance Indicators (OPIs)—can meaningfully reflect variations in technical execution and correlate with postoperative recovery in robotic colorectal cancer resections. A large number of task-specific OPIs were found to significantly associate with hospital length of stay and patient-reported condition scores. These correlations were especially pronounced during technically demanding phases such as transection and vessel ligation, suggesting that precise motor control and efficient instrument handling during these steps are critical to short-term outcomes. Furthermore, machine learning models trained on selected OPIs achieved moderate accuracy in predicting both outcome measures, validating the potential of OPIs as inputs for real-time performance monitoring systems.

Despite traditional assumptions, conventional surgeon experience indicators (e.g., years in practice or annual case volume) were not predictive of patient outcomes in this cohort. This underscores the need for granular, behavior-based evaluation frameworks over generalized experience proxies. Finally, the development of surgeon-specific feedback reports provided a novel way of translating data-driven analysis into actionable insights, supporting personalized surgical education.

5.2 Implications for Surgical Education

These findings offer several implications for surgical training. First, OPIs enable the objective quantification of micro-behaviors that are often missed by human evaluators. For instance, excessive wrist articulation or prolonged tool activation—found to correlate with poorer outcomes—can be highlighted in structured feedback. Training curricula can incorporate this feedback to emphasize the modulation of specific instrument usage patterns and ergonomic strategies. Second, personalized performance reports facilitate targeted remediation and progression tracking, especially for trainees transitioning to independent robotic practice. Rather than relying solely on subjective assessments, educators can use benchmarked OPI distributions to guide coaching interventions grounded in empirical associations with patient outcomes.

5.3 Integration into Clinical Practice

The integration of OPI tracking into real-world operating rooms is technically feasible, given that robotic systems already generate high-resolution console logs. A future clinical workflow might include real-time OPI capture, automatic risk stratification, and post-case performance summaries. These could be reviewed during surgical debriefings, akin to pilot cockpit logs. Repeated OPI assessment across a surgeon's case portfolio could inform recertification protocols or continuous professional development, shifting credentialing from volume-based to competency-based models. The interpretability and automation of the feedback system developed in this work positions it as a candidate for integration into existing robotic platforms, supporting both surgeon self-assessment and institutional quality assurance initiatives.

5.4 Limitations

This study has several limitations. First, the dataset is limited in size, both in terms of patient operations (n=41) and surgeon diversity. A single surgeon accounted for more than half of all cases, potentially biasing task distributions and metric variance. Second, while the OPI set used was extensive, it remains constrained to the metrics provided by the da Vinci Xi's logging system, excluding potentially important variables such as instrument torque or tissue force, which are not routinely recorded. Third, the results may not generalize beyond colorectal procedures or to laparoscopic/open contexts. Lastly, the PCA-derived patient condition score is a composite measure and lacks a direct clinical interpretation, though it captures meaningful variance in recovery-related symptoms.

Chapter 6

Future Work

6.1 Expansion to Other Surgical Specialties & Procedures

While this study focused on robotic colorectal cancer resections, the OPI framework can be extended to other domains of robot-assisted surgery. Urologic and gynecologic oncology procedures, such as radical prostatectomy or hysterectomy, also require precise pelvic dissection and could benefit from similar skill assessment paradigms. Furthermore, multicenter studies with harmonized OPI definitions would allow for larger sample sizes, diverse surgeon profiles, and broader validation of task-outcome relationships. Such collaborations could form the foundation for robust, cross-specialty performance benchmarks.

6.2 Computer Vision for Automated Task Annotation

Currently, task segmentation is performed manually or extracted from annotations, which limits scalability. Future work could integrate computer vision techniques to automate task recognition using intraoperative video feeds. Deep learning models trained on annotated surgical videos—such as convolutional neural networks (CNNs)—could identify phases of the procedure in real time, thereby enabling fully automated OPI extraction per task. Combining video-derived kinematics (e.g., tool trajectory, orientation, and collisions) with robotic console metrics could yield more comprehensive assessments of technical skill and enable live feedback.

6.3 Enhancing Feedback Interpretability

While the current feedback report offers percentile-based insights, its interpretability could be further enhanced by including visualizations, temporal comparisons, and suggested improvement strategies. For example, graphical overlays of a surgeon's metric trajectory over time could help monitor skill progression. Further, incorporating surgeon responses or reflections into the report might promote reflective practice and continuous improvement, reinforcing the learning cycle.

Chapter 7

Conclusion

7.1 Summary of Contributions

This project systematically demonstrated that robotic console-derived metrics, or Operative Performance Indicators (OPIs), can be used to assess surgeon performance and predict postoperative outcomes in robotic colorectal cancer surgery. A novel contribution was the identification of task-specific OPI-outcome correlations across hospital length of stay and patient-reported condition scores. These insights were operationalized into a personalized surgeon feedback system, representing a practical step toward data-driven skill refinement. Additionally, predictive models trained on these metrics confirmed their explanatory power and hinted at the feasibility of outcome forecasting from intraoperative behavior alone.

7.2 Final Remarks on OPIs in Robotic Colorectal Cancer Resection

As robotic surgery becomes increasingly prevalent, the need for scalable, objective, and meaningful performance assessment tools grows in parallel. This work establishes a technical and conceptual framework for using OPIs to bridge that gap. By aligning skill metrics with clinical outcomes and embedding them into interpretable reports, this approach supports a shift from experience-based to evidence-based surgical education. Wider adoption of such methods could lead to more personalized training, more precise credentialing, and ultimately, better patient outcomes through continuous refinement of surgical technique.

Bibliography

[1] B. W. Vabi, J. F. Gibbs, and G. S. Parker. Implications of the growing incidence of global colorectal cancer. *Journal of Gastrointestinal Oncology*, 12(Suppl 2): S387–S398, 2021. doi: 10.21037/jgo-2019-gi-06.

[2] J. Franklyn, M. Abdalkoddus, B. Rossi, D. Limnatitou, M. Ahmed, A. Sokker, W. J. Douie, and S. Smolarek. Prognosis of non-operative management of non-metastatic colorectal cancer in octa- and nonagenarians. *Annals of the Royal College of Surgeons of England*, 102(7):504–509, 2020. doi: 10.1308/rcsann.2020.0067.

[3] B. C. Gulack, D. P. Nussbaum, J. E. Keenan, A. M. Ganapathi, Z. Sun, M. Worni, J. Migaly, and C. R. Mantyh. Surgical resection of the primary tumor in stage iv colorectal cancer without metastasectomy is associated with improved overall survival compared with chemotherapy/radiation therapy alone. *Diseases of the Colon Rectum*, 59(4):299–305, 2016. doi: 10.1097/DCR.0000000000000546.

[4] F. Polat, L. H. Willems, K. Dogan, and C. Rosman. The oncological and surgical safety of robot-assisted surgery in colorectal cancer: outcomes of a longitudinal prospective cohort study. *Surgical Endoscopy*, 33(11):3644–3655, 2019. doi: 10.1007/s00464-018-06653-2.

[5] S. Park and N. K. Kim. The role of robotic surgery for rectal cancer: Overcoming technical challenges in laparoscopic surgery by advanced techniques. *Journal of Korean Medical Science*, 30(7):837–846, 2015. doi: 10.3346/jkms.2015.30.7.837.

[6] S. H. Baik, C. M. Kang, W. J. Lee, N. K. Kim, S. K. Sohn, H. S. Chi, and C. H. Cho. Robotic total mesorectal excision for the treatment of rectal cancer. *Journal of Robotic Surgery*, 1(1):99–102, 2007. doi: 10.1007/s11701-007-0015-0.

[7] A. Pai, G. Melich, S. J. Marecik, J. J. Park, and L. M. Prasad. Current status of robotic surgery for rectal cancer: A bird's eye view. *Journal of Minimal Access Surgery*, 11(1):29–34, 2015. doi: 10.4103/0972-9941.147682.

[8] A. F. Adili, J. Di Giovanni, E. Kolesar, N. C. Wong, J. Hoogenes, S. Dason, and B. Shayegan. Positive surgical margin rates during the robot-assisted laparoscopic radical prostatectomy learning curve of an experienced laparoscopic surgeon. *Canadian Urological Association Journal*, 11(11):E409–E413, 2017. doi: 10.5489/cuaj.4588.

[9] C. Y. Lin, Y. C. Liu, M. C. Chen, and F. F. Chiang. Learning curve and surgical

outcome of robotic assisted colorectal surgery with eras program. *Scientific Reports*, 12(1):20566, 2022. doi: 10.1038/s41598-022-24665-w.

[10] Daniel S. Oh, Marzieh Ershad, Jon O. Wee, Manu S. Sancheti, Desmond M. D'Souza, Luis J. Herrera, Lana Y. Schumacher, Mallory Shields, Kristen Brown, Sadia Yousaf, and John F. Lazar. Comparison of global evaluative assessment of robotic surgery with objective performance indicators for the assessment of skill during robotic-assisted thoracic surgery. *Surgery*, 174(6):1349–1355, 2023. doi: 10.1016/j.surg.2023.08.008.

[11] Jose T. Larach and Satish K. Warrier. Robotics and transanal total mesorectal excision. *Annals of Laparoscopic and Endoscopic Surgery*, 5(0), 2020. ISSN 2518-6973. URL `https://ales.amegroups.org/article/view/5777`.

[12] William D. Gerull, Sue Kulason, Mallory C. Shields, Andrew Yee, and Michael M. Awad. Impact of robotic surgery objective performance indicators: A systematic review. *Journal of the American College of Surgeons*, 240(2):201–210, 2025. doi: 10.1097/XCS.0000000000001208.

[13] C. W. Huang, H. L. Tsai, Y. S. Yeh, and et al. Robotic-assisted total mesorectal excision with the single-docking technique for patients with rectal cancer. *BMC Surgery*, 17:126, 2017. doi: 10.1186/s12893-017-0315-x.

[14] Francesco Marchegiani, Chiara A. Schena, G. Santambrogio, E. P. Emma, I. Tsimailo, and N. de'Angelis. Total mesorectal excision with new robotic platforms: A scoping review. *Journal of Clinical Medicine*, 13:6403, 2024. doi: 10.3390/jcm13216403.

[15] Andrew J. Hung, Jian Chen, Zhengping Che, Tanachat Nilanon, Anthony Jarc, Micha Titus, Paul J. Oh, Inderbir S. Gill, and Yan Liu. Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. *Journal of Endourology*, 32(5):438–444, 2018.

[16] Andrew J. Hung, Paul J. Oh, Jian Chen, Seyed Ghodoussipour, Christopher Lane, Anthony Jarc, and Inderbir S. Gill. Experts vs super-experts: differences in automated performance metrics and clinical outcomes for robot-assisted radical prostatectomy. *BJU International*, 123:861–868, 2019. doi: 10.1111/bju.14599.

[17] William D. Gerull, Sue Kulason, Mallory C. Shields, Andrew Yee, and Michael M. Awad. Impact of robotic surgery objective performance indicators: A systematic review. *Journal of the American College of Surgeons*, 240(2):201–210, 2025. doi: 10.1097/XCS.0000000000001208.

[18] Florian Neis, Sara Y. Brucker, Alexander Bauer, Mallory Shields, Laura Purvis, Xiaoyu Liu, Mohammad Ershad, Charlotte B. Walter, Tim Dijkstra, Christoph Reisenauer, and Bernd Kraemer. Novel workflow analysis of robot-assisted hysterectomy through objective performance indicators: a pilot study. *Frontiers in Medicine*, 11:1382609, 2024. doi: 10.3389/fmed.2024.1382609.

[19] Royal College of Surgeons. Measuring the quality of surgical care

and setting benchmarks for training using intuitive data recorder technology. https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/mastery/, 2020. Accessed 6 Apr 2025. Sponsored by the University of Birmingham. REC reference: 20/PR/0580.

[20] Intuitive Surgical. Da vinci xi robotic surgical system, 2025. URL https://www.intuitive.com/en-us/products-and-services/da-vinci/xi. Accessed: 2025-04-06.

[21] Intuitive Foundation. Robotic surgery fellowship funding, 2025. URL https://www.intuitive-foundation.org/fellowships/. Accessed: 2025-04-06.

Appendix A

Abbreviations

Abbreviation	Meaning
BMI	Body Mass Index
CNN	Convolutional Neural Network
GEARS	Global Evaluative Assessment of Robotic Skills
GOALS	Global Operative Assessment of Laparoscopic Skills
IDR	Intuitive Data Recorder
LOS	Length of Stay
MAE	Mean Absolute Error
MASTERY	Metrics and Analytics to Support Training and Evaluation of Robotic sYstems
MISTELS	McGill Inanimate System for Training and Evaluation of Laparoscopic Skills
OPI	Operative Performance Indicator
OSATS	Objective Structured Assessment of Technical Skills
PCA	Principal Component Analysis
PC1	Principal Component 1
RMSE	Root Mean Squared Error
SSC	Surgeon Side Console
SVR	Support Vector Regression
TME	Total Mesorectal Excision
USM	User Side Manipulator
XGBoost	eXtreme Gradient Boosting