Developing a Multiparametric OPI-based Proficiency Score for Robotic Inguinal Hernia Repair

Mattia Ballo[1,2], Emaad Iqbal[1], Maria Clara Morais[1], Aditya Amit Godbole[1], Andrew Yee[3], Filippo Filicori[1]

1) Intraoperative Performance Analytics Laboratory, Department of Surgery, Lenox Hill Hospital, New York, NY, USA; 2) Department of Surgical, Oncological and Gastroenterological Sciences, General Surgery, Hepato-Pancreato-Biliary Surgery and Liver Transplantation, Padua University Hospital, PD, Italy. , 3) Intuitive Surgical, Atlanta, GA, USA

Introduction

Robotic-assisted surgery is integral to the the training of General Surgery residents. There is a need for curriculum building and tools for assessment of performance. Objective performance indicators (OPIs) derived from robotic system kinematic data can be used as quantitative measures of trainee proficiency. Building on evidence that OPIs improve with experience, our primary objective was to develop and characterize a continuous, multiparametric OPI-based performance score that quantifies intraoperative surgeon performance during key tasks of robotic inguinal hernia repair (IHR) and explore the association of this score with trainee progression

Methods

Robotic inguinal hernia repairs (n=277) with kinematic data from Intuitive Data Recorder between October 2023 and September 2025 at a single institution were included. Analysis focused on two critical procedural steps: (1) creation of the peritoneal flap and (2) closure of the peritoneum. Phase was assigned to trainee group if trainee console was active >75% of the phase time.

From an initial panel of 50 console-derived OPIs, 17 were retained after multicollinearity screening. An ordinal regression classifier (LogisticAT, mord) was trained with three target groups (PGY3–4, PGY5–6, attending). Attending cases were stabilized for model fitting by retaining only attending cases with operative duration within the interquartile range and performed beyond each surgeon's median case number.

Model probability outputs were transformed into a continuous 0–100 performance score and analyzed over time. Linear regression assessed correlation of these scores with time since beginning of training.

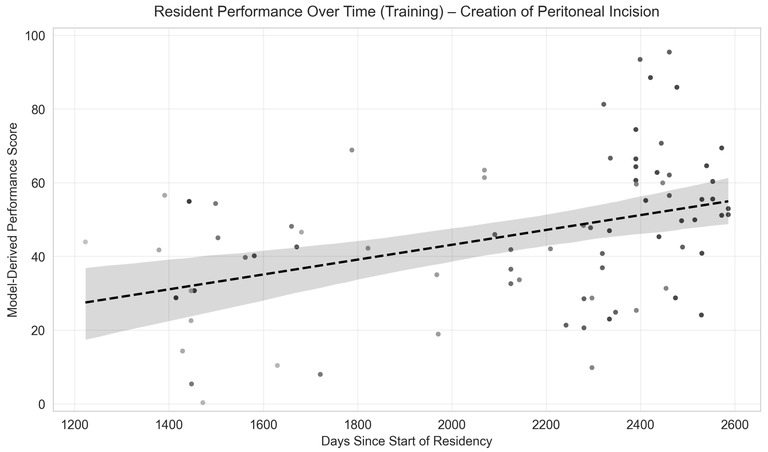

Figure 1: Resident performance at creation of peritoneal flap over time since beginning of training.

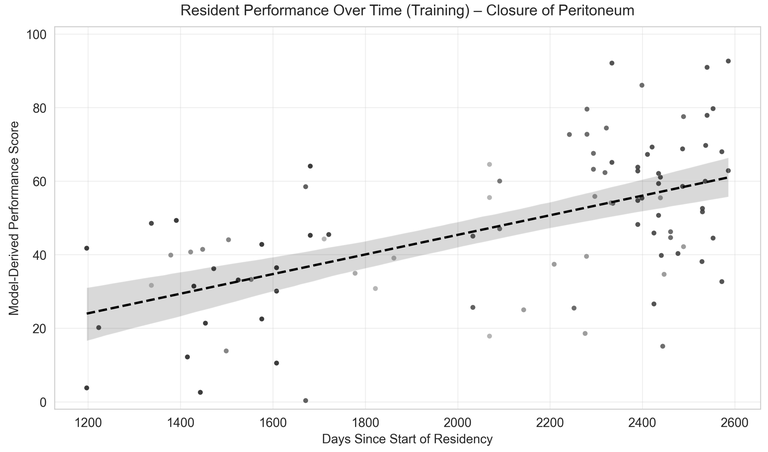

Figure 2: Resident performance at closure of peritoneal flap incision over time since beginning of training.

Results

For peritoneal incision (456 segments), the model achieved 71.5% accuracy, 57.0% balanced accuracy, and weighted F1 of 0.77 in predicting the surgeon's experience group. For peritoneal closure (446 segments), accuracy was 65.7%, balanced accuracy 58.5%, and weighted F1 0.68. The model-derived performance score correlated significantly with surgical experience: in learning-curve analysis, the score for peritoneal incision (Figure 1) demonstrated a positive association with progression ($R = 0.26$, $p = 0.01$). and a significant association was observed for peritoneal closure ($R = 0.48$, $p < 0.001$, Figure 2) .

Discussion

OPI-based modeling provides an objective and scalable framework to evaluate surgical proficiency in robotic procedures. We developed an OPI-based performance score that differentiates expertise levels and captures progression across training. Unlike traditional assessments, this approach objectively quantifies console performance continuously and supports automated feedback, offering a pathway toward competency-based data driven surgical education. Further work will focus on actionability of the insights offered and expansion of the model to other phases of IHR and other procedures.

Conclusion

OPI based performance scoring algorithms allow for objective means to measure trainee performance. These have significant potential as general surgery training programs move to Entrustable Professional Activities for trainee evaluation.