Video replay in surgery: Can we make the "right call" in predicting outcomes?

Edward D. Matsumoto, MD, MEd, FRCSC

Division of Urology, McMaster University, Hamilton, ON, Canada

In their article, Goldenberg et al set out to answer an important question in surgical education: Can we assess technical performance in the operating room to predict clinical outcomes?

The authors compared assessments by expert raters (robotic and open surgeons) and those from lay people from the Crowd-Sourced Assessment of Technical Skills (C-SATS)[1,2] group with regard to experienced surgeons' technical ability during a robotic-assisted radical cystectomy. The authors prepared short video segments (60 seconds) showing mobilization of the ureter, ureteral preparation for the anastomosis, and the ureteral-ileal anastomosis from nine (out of a potential 102) cases that resulted in clinically significant postoperative uretero-ileal strictures (UIS) (10 strictures in total). They compared these to video segments showing the same steps from eight control cases that did not result in UIS. Of note, the control group consisted of the same patients, but the video was of the procedure on the contralateral ureter that did not develop a stricture. Five content experts rated each step on a five-item dichotomous (yes/no) questionnaire developed for this study. The questionnaire assessed perceived risk for UIS during each step and overall. The C-SATS group (2142 lay people) used the Global Evaluative Assessment of Robotic Skills (GEARS) global rating scale to assess the videos. Both the authors and the C-SATS workers found there was no association between the experts' scores and UIS.

The authors have taken on an ambitious task of trying to predict a clinical outcome based on surgical performance. Great gains have been made in technical skills assessment; however, the majority of the literature has reported on surgical performance in the ex-vivo operative environment (i.e., surgical simulation laboratories). To date, largely due to ethical issues, there has been a paucity of studies that evaluate intraoperative assessment. As academic surgeons in urology, we owe it to our trainees, and more importantly to our patients, to further improve surgical performance and ultimately clinical outcomes.

Limitations of this study were clearly pointed out: the small sample size and that the non-validated assessment tool and GEARS may not be sensitive enough to identify the nuances of the specific steps identified as being critical in avoiding UIS. Developing a tool (or even adapting a previously validated metric) to capture the nuances of a surgical step is paramount in seeking a valid and reliable intraoperative assessment.

I also suspect, as we start assessing more complex procedures, it will become more inherent that the raters are properly trained to use the given assessment tool. The notion of being able to incorporate crowd-sourcing into assessment of intraoperative videos of residents at different levels of training, or even staff still in the learning curve of a specific type of case, is very appealing. Crowd-sourcing may be suitable for assessing basic skills, such a laparoscopic skills, and sensitive enough to discern between novice and expert performance;[1] however, I postulate that in assessment, especially when it comes to high-stakes assessment, an expert's knowledge of the procedure and understanding of the nuances will play a critical role.

Focusing ongoing research in assessment of surgical performance in the high-stakes operating room environment is a direction we as academics must seek. There is much to be desired when it comes to the development, validation, and use of rating tools and the selection and training of assessors. Moving forward, I believe we will be seeing important developments in this area, as medical education continues to evolve toward meeting greater societal expectations and assessing trainees' surgical competency.

References

1. Kowalewski TM, Comstock B, Sweet R, et al. Crowd-sourced assessment of technical skills for validation of basic laparoscopic urologic skills (BLUS) tasks. *JURO* 2016:1-15. https://doi.org/10.1016/j.juro.2016.01.005
2. Ghani KR, Miller DC, Linsell S, et al; the Michigan Urological Surgery Improvement Collaborative. Measuring to improve: Peer and crowd-sourced assessments of technical skill with robot-assisted radical prostatectomy. *Eur Urol* 2016;69:547-50. https://doi.org/10.1016/j.eururo.2015.11.028

Correspondence: Dr. Edward D. Matsumoto, Division of Urology, McMaster University, Hamilton, ON, Canada; matsumo@mcmaster.ca