



ASSESSMENT OF SURGICAL SKILLS OF UROLOGY RESIDENT APPLICANTS: FOUR-YEAR EXPERIENCE

Rajiv Karani, Francis A. Jefferson, Krista N. Larson, Simone Vernez, Lillian Xie, Roshan M. Patel, Kathryn Osann, Shlomi Tapiero, Zhamshid Okhunov, Heidi A. Stephany, Jaime Landman, Ralph V. Clayman

Department of Urology, University of California, Irvine, USA



INTRODUCTION

Urology residents are often chosen based on USMLE scores, interview performance, and letters of recommendation. Surgical skills testing is notably absent. We sought to determine if crowd-sourcing can effectively assess surgical skills among resident applicants and to determine where skills testing ranks in relation to other highly valued applicant metrics.

METHODS

- Ninety-four urology residency interviewees between 2016-2019 were recorded performing knot tying, laparoscopic peg transfer, and robotic suturing (**Figure 1**).
- Videos were uploaded to the Crowd-Sourced Assessment of Technical Skills (C-SATS) website for evaluation. C-SATS scoring was compared to scoring by two faculty urologists.
- Validated tools were used for scoring: Objective Structured Assessment of Technical Skills (OSATS) for knot tying, Global Operative Assessment of Laparoscopic Skills (GOALS) for laparoscopic peg transfer, and Global Evaluative Assessment of Robotic Skills (GEARS) for robotic suturing.
- USMLE Step 1 scores, interview scores, and match rank were recorded.
- Follow-up surveys were administered regarding the interviewees' match results and career plans.
- Agreement between crowd and faculty was evaluated using Cronbach's alpha (α). Pearson correlation testing determined the relationships among the applicant metrics.

RESULTS

- Faculty and C-SATS scoring showed good agreement ($\alpha = 0.88$) only for laparoscopic peg transfer (Table 1).
- Crowd: OSATS, GOALS, and GEARS correlated poorly with final match rank (Table 2).
- Faculty: OSATS and GEARS correlated poorly with final match rank (Table 2).
- Among all recorded metrics, interview score most closely correlated with final match rank ($r = -0.72$, $p < 0.001$) (Table 2).
- On a general linear model, interview score and GOALS faculty score were predictive of final match rank.
- Interviewees who reported matching at one of their top 3 choices had significantly higher GOALS faculty scores (11.9 vs. 9.7, $p = 0.03$).
- The majority of participants reported future plans of a surgical-based practice.

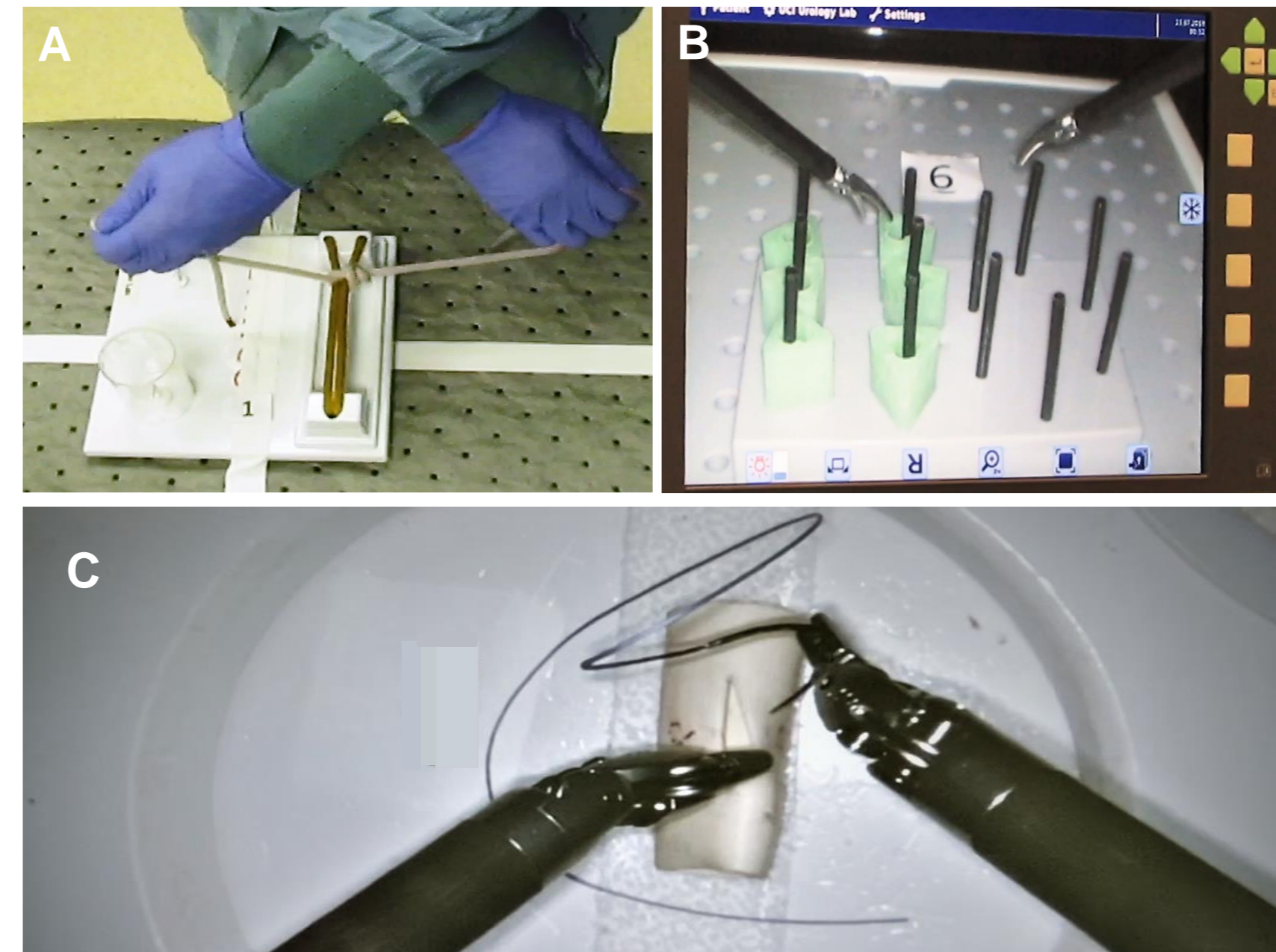


Figure 1. A) Open knot tying B) Laparoscopic peg transfer C) Robotic suturing

Table 1. Agreement Between C-SATS and Faculty Scores

	Mean	Cronbach Alpha Agreement
OSATS Crowd	16.7 ± 1.0	0.32
OSATS Faculty	12.3 ± 3.8	
GOALS Crowd	10.8 ± 2.9	0.88
GOALS Faculty	11.4 ± 3.8	
GEARS Crowd	15.5 ± 1.7	0.66
GEARS Faculty	11.6 ± 4.8	
Average Score for all 3 Tasks Crowd	13.3 ± 1.4	0.73
Average Score for all 3 Tasks Faculty	11.0 ± 2.9	

Cronbach alpha scale of agreement: >0.9 = excellent; 0.899-0.7 = good; 0.699-0.5 = fair; <0.5 = poor agreement

Table 2. Correlation Coefficients (r) Between Applicant Metrics

	Correlation with Match Rank	p-value
USMLE Step 1	-0.14	0.25
Interview Score	-0.72	<0.001
Average Skills Test Score Crowd	-0.14	0.25
OSATS Crowd	0.03	0.78
GOALS Crowd	-0.09	0.44
GEARS Crowd	-0.15	0.18
Average Skills Test Score Faculty	-0.12	0.32
OSATS Faculty	0.007	0.95
GOALS Faculty	-0.15	0.21
GEARS Faculty	-0.13	0.26

CONCLUSIONS

- Expert faculty, but not crowd-sourced, assessment of surgical skills among urology resident applicants aids in determining match rank.
- Applicants who perform better on laparoscopic peg transfer are more likely to match at a top 3 choice.
- Interview performance and faculty scored laparoscopic peg transfer are associated with final match rank.