Engineering and Urology Society

32\textsuperscript{nd} Annual Meeting

Friday May 12\textsuperscript{th}, 2017

Boston, MA

http://engineering-urology.org/
ABSTRACT 18

HIGH VOLUME SKILL ASSESSMENT FOR BASIC LAPAROSCOPIC UROLOGIC SKILLS (BLUS)

Timothy M Kowalewski,1 Bryan Comstock,2 Robert Sweet,3 Cory Schaffhausen,1 Ashleigh Menhadji,4 Timothy Averch,5 Geoffrey Box,6 Timothy Brand,7 Michael Ferrandino,8 Jihad Kaouk,9 Bodo Knudsen,6 Jaime Landman,10 Benjamin R. Lee,11 Bradley F. Schwartz,12 Elspeth McDougall13 and Thomas S. Lendvay14

1 University of Minnesota Mech. Engineering,
2 University of Washington Biostatistics,
3 University of Washington Urology,
4 Boston University School of Medicine,
5 University of Pittsburgh Medical Center
6 Ohio State University Urology
7 Madigan Army Medical Center, USUHS
8 Duke University, Urology
9 Cleveland Clinic, Urology
10 University of California Irvine, Urology
11 University of Arizona, Urology
12 Southern Illinois University, Urology
13 University of British Columbia, Urologic Sciences,
14 University of Washington Urology and Seattle Children’s Hospital

Introduction: In an effort to explore a scalable and rapidly workable methodology for large-scale assessments for Basic Laparoscopic Urologic Skills (BLUS) [PMID22050489, 25911459], a consortium of eight urologic laparoscopic training centers evaluated concurrent validity [PMID26527514] by comparing skills assessment data from crowdsourcing [PMID26421369] with both expert evaluations [PMID26778711] and computerized skill measurement methodologies[PMID20607563, 25108691].

Methods: Surgical performance videos (N=430) of BLUS tasks (Peg Transfer, Cutting, Suturing, and Clip Apply) were subjected to anonymized crowdsourced ratings using the same surveys tools as expert rater panels: Global Objective Assessment of Laparoscopic Skills (GOALS) [PMID15972181] ratings from a faculty panel (N=5). Prerecorded tool motion metrics (tool path length, jerk cost, grasp forces, etc.) were available for each video using the Electronic Data Generation and Evaluation (EDGE) drylab box trainer (Simulab, WA). We evaluated concordance between mean crowd scores to both motion metrics and expert ratings. We used Cronbach’s alpha, Spearman’s rho (>0.8 indicates very strong agreement), and Receiver Operator Curves (ROC) with Area Under Curve (AUC) statistics. We considered correlations statistically significant only at p < 0.01.

Results: Crowdsowers provided 16,418 individual reviews of 430 videos in 8.7 days. Spearman’s rho between composite crowd scores and motion metrics ranged from moderately strong to very strong (Peg Transfer 0.84, Cutting 0.72, Suturing, 0.80 and Clip Apply 0.61). ROC curves for each BLUS task (AUC=0.93 Peg Transfer, 0.93 Cutting, 0.88 Suturing, and 0.79 Clip Applying) was consistent with expert faculty skills assessment. GOALS ratings of expert panels and crowds had moderate to strong concordance (0.70 to 0.95) for a range of recordings from novices to experts.

Conclusion: High volume assessment via crowd-sourcing is feasible and concordant with objective BLUS psychomotor tool motion metrics. These surgical skill assessments align with those of expert faculty. Crowdsourcing offers a rapid, scalable approach to the determina-tion of learner proficiency.

Figure: (Left) The EDGE motion capture system used to record tool motion: x,y,z, roll, grasp angle and force and (Right) Representative ROC curve of Crowd Ratings vs EDGE motion metrics cutting (Area Under Curve shown; desired perfect agreement at AUC=1).