Evaluating the Evaluators: Does C-SATS Measure Up?

Robert B Laverty, MD[1]; Charles H Chesnut, MD[2]; Joseph R Karam, MD[3]; Joseph L'Huillier, MD[4]; Alexander Bonte, MD[5]; Julie Clanahan, MD[6]; Robert W Krell, MD[1]

[1]Department of Surgery, Brooke Army Medical Center, San Antonio, TX [2]General Surgery Residency – Valley System Health, Las Vegas, NV [3] Start Center for Surgical Care, San Antonio, TX [4] Department of Surgery, University of Buffalo, Buffalo, NY [5] Department of Surgery, Hackensack University Medical School, Hackensack, NJ [6] Department of Surgery, Washington University, St Louis, MO

Background

- Robotic-assisted surgery has increased in prevalence, particularly in general surgery.
- The number of cases required to achieve adequate proficiency in robotic surgery, however, and the training metrics that correlate best with proficiency, remain unclear.
- We sought to better define proficiency-based benchmarks in robotic-assisted (RA) cholecystectomies and inguinal hernia repairs (IHR) using Crowd-Sourced Assessment of Technical Skill (C-SATS), a company that markets its ability evaluate surgical competency.

Methods

- Multi-institutional, cross-sectional study in which 48 surgeons (senior residents, fellows and practicing physicians) submitted representative videos of themselves performing a RA cholecystectomy and/or IHR that subsequently underwent a blinded review using the C-SATS platform, which utilizes the Global Evaluative Assessment of Robotic Skills (GEARS) rubric.
- Primary outcome was correlation of GEARS scores with historic procedure case volume. Secondary outcomes included construct validity of GEARS scores as an operative proficiency metric.

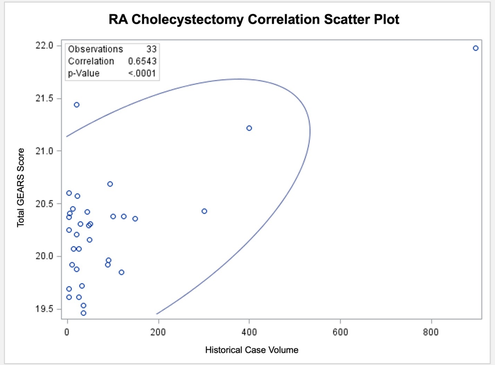
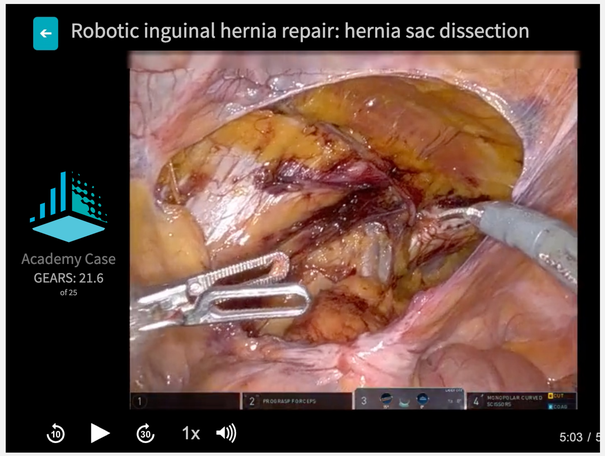
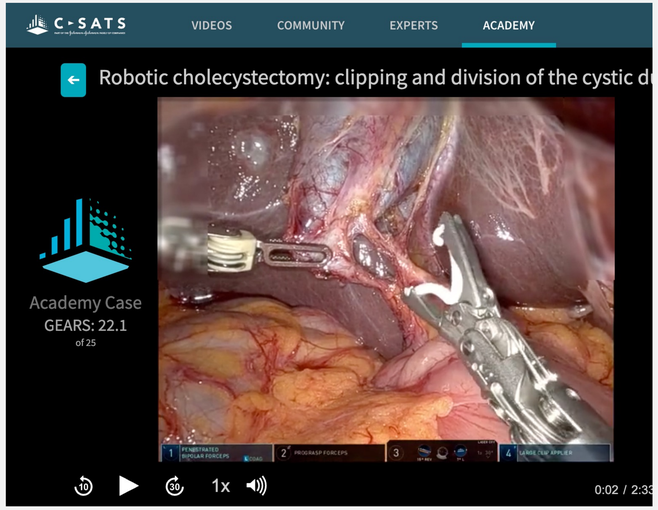


Figure 1A. Correlation between total GEARS scores and historical case volume of all surgeons performing RA cholecystectomy, scatter plot with 95% prediction ellipses.





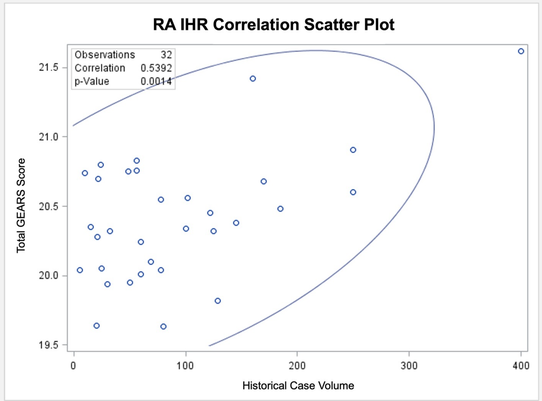


Figure 1B. Correlation between total GEARS scores and historical case volume of all surgeons performing RA IHR, scatter plot with 95% prediction ellipses.

Results

- Significant correlations were demonstrated between total GEARS scores and historical case volume for both RA cholecystectomy (r=0.65, p<0.0001) and IHR (r=0.54, p=0.001) amongst all performers (**Figure 1A/B**), however, statistical significance was not achieved for correlations amongst resident/fellow physicians (r=0.39, p=0.11 for cholecystectomy; r=0.22, p=0.49 for IHR) or those with <50 historic case volume (r=0.14, p=0.55 for cholecystectomy; r=0.21, p=0.54 for IHR).
- No difference in total GEARS scores was seen between resident/fellow and practicing physicians for either RA cholecystectomy (20.21 v 20.25, p=0.82) or RA IHR (20.45 v 20.46, p=0.95), nor in those with <50 or ≥50 historic case volume in RA cholecystectomy (20.16 v 20.33, p=0.33) and RA IHR (20.35 v 20.49, p=0.48).
- GEARS scores by domain (bimanual dexterity, depth perception, efficiency, force sensitivity and robotic control) and surgical step (exposure of triangle of calot, clipping and division of cystic artery/duct, and dissection of gallbladder; mobilizing peritoneal flap, hernia sac dissection and mesh placement) were similar across both groups (p>0.05).

Conclusions

- C-SATS-derived GEARS scores correlated to overall surgeon historical case volume for RA cholecystectomy and IHR, but not amongst novice performers. This methodology was unable to differentiate between novice and expert performers in our surgeon population calling into question its utility as an instrument to assess technical surgical skill.

Disclaimer: The view(s) expressed herein are those of the author(s) and do not reflect the official policy or position of Brooke Army Medical Center, the US Army Medical Department, the Department of the Army, Department of the Air Force, Department of Defense, or the US Government. The voluntary, fully informed consent of the subjects used in this research was obtained as required by 32 CFR 219 and DODI 3216.02_AFI40-402.