# Video-based assessment for laparoscopic fundoplication: Validity evidence for rater training among intermediate-level surgical trainees

SAGES VBA for LF Rubric

Dylan J Grisell MD[1] & Noosha D Deravi MD[2], Aimee K Gardner PhD[3,8], Marin S Kurian MD[4], Anna R. Ibele MD[5], Nova Szoka MD[6], Linda Schultz[7], Liane S Feldman MD[1], E. Matthew Ritter MD MPHE[2]

*1 – Dept of Surgery, McGill University, Montreal, QC, CA; 2 – Dept of Surgery, Indiana University School of Medicine, Indianapolis, IN; 3 – Dept of Surgery, University of Colorado Anschutz School of Medicine, Denver, CO; 4 – Dept of Surgery, NYU Langone Health, New York, NY; 5 – Dept of Surgery, University of Utah School of Medicine, Salt Lake City, UT; 6 – Dept of Surgery, West Virginia University School of Medicine, Morgantown, WV; 7 – Society of American Gastrointestinal and Endoscopic Surgeons, Los Angeles, CA; 8 – SurgWise LLC, Denver, CO*

## Introduction

- Surgical skill correlates with patient outcomes[1-6]
- There is a lack of procedure-specific tools to evaluate performance by video-based assessment (VBA)[3-5]
- The SAGES VBA for laparoscopic fundoplication (LF) rubric (QR code above) was developed through a rigorous process involving rater training of expert foregut surgeons using eLearning modules[3]
- Relying solely on expert assessors may challenge scalability

**Objective** – To generate pilot data to examine the validity of the VBA for LF rater training system through its use by non-expert assessors

## Methods

- Participants: Two PGY3 General Surgery residents
- Rater Training: Series of online learning modules describing the components of the VBA for LF rubric supplemented with practice assessments using short video segments to standardize scoring
- Video Review: Scoring using the VBA for LF rubric
  1) Round 1: 10 full-length LF videos
  2) Disagreement Resolution: Residents met with SMEs to discuss scoring discrepancies and strengthen a shared mental model
  3) Round 2: 8 full-length LF videos
- Data Analysis: Percent agreement (PA) for procedure sub-steps, mean PA by video, and pass/fail (using a pre-established cut score) were calculated between rater pairs for each round, then compared to expert data from the initial rater training study

## Results

- Procedure sub-steps with <75% PA after Round 1 included components of:
  - Hiatal dissection – Quality
  - Esophageal Mobilization – Skill
  - Hiatal Repair – Quality
- Mean PA for all rater pairs increased from Round 1 to Round 2 after disagreement resolution meeting

| | Resident A Resident B | Resident A Expert A | Resident B Expert A | Experts |
|---|---|---|---|---|
| Round 1 | 80% | 85% | 84% | 75% |
| Round 2 | 95% | 95% | 98% | 91% |

Table 1: Mean percent agreement (PA)

- Mean pass rate among residents was significantly lower than experts in Round 1, but groups were comparable in Round 2 (75% vs. 70.8%)

| | Resident A | Resident B | Expert A | Expert B | Expert C |
|---|---|---|---|---|---|
| Round 1 | 10% | 20% | 80% | 50% | 0% |
| Round 2 | 62.5% | 87.5% | 100% | 37.5% | 75% |

Table 2: Pass rates by individual rater

| | Residents | Experts |
|---|---|---|
| Round 1 | 15% | 43.3% |
| Round 2 | 75% | 70.8% |

Table 3: Mean pass rates by rater group

- PA for pass/fail between residents and experts increased from 30% (3/10 videos) in Round 1 to 62.5% (5/8 videos) in Round 2

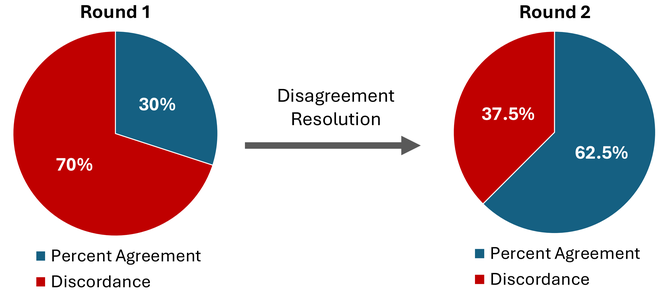



Figure 1: PA and discordance on pass/fail between residents and experts

## Conclusions

- Pilot data demonstrates application of the SAGES VBA for LF rater training tool among non-expert assessors
- Increases in inter-rater reliability after the disagreement resolution meeting highlights areas for improvement in the rater training process
- Future, larger studies will be required to contribute additional validity evidence for use of the VBA for LF in surgical competency

## References

1. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med*. 2013;369(15):1434-1442
2. Balvardi S, Kammili A, Hanson M, et al. The association between video-based assessment of intraoperative technical performance and patient outcomes : A systematic review. *Surg Endosc*. 2022;36(11):7938-7948.
3. Ritter EM, Gardner AK, Dunkin BJ, et al. Video-based assessment for laparoscopic fundoplication: Initial development of a robust tool for operative performance assessment. *Surg Endosc*. 2019;34(7):3176-3183.
4. Balvardi S, Semsar-Kazerooni K, Kaneva P, et al. Validity of video-based general and procedure-specific self-assessment tools for surgical trainees in laparoscopic cholecystectomy. *Surg Endosc*. 2022;37(3):2281-9.
5. Feldman LS, Pryor AD, Gardner AK, et al. SAGES Video-Based Assessment (VBA) program: A vision for life-long learning for surgeons. *Surg Endosc*. 2020;34(8):3285-8.
6. Ryan JF, Mador B, Lai K, et al. Validity evidence for procedure-specific competence assessment tools in general surgery. *Ann Surg*. 2022;275(3):482-7.