

ORIGINAL RESEARCH

Crowd-Sourced Assessment of Laparoscopic Surgical Skills for Ob/Gyn Residents

Bayley Clarke, M.D.^{1,2} , Jovana Tavcar, M.D.^{2,3}, Nicholas Hazen, M.D.^{2,3}, James Robinson, M.D.⁴

¹ Obstetrics and Gynecology, MedStar Washington Hospital Center, ² Obstetrics and Gynecology, MedStar Georgetown University Hospital, ³ Georgetown University School of Medicine, ⁴ Minimally Invasive Gynecology Surgery, Shady Grove Fertility Center

Keywords: feedback, graduate medical education, resident evaluation, crowd-sourcing, surgical education, laparoscopy

<https://doi.org/10.52504/001c.124417>

Georgetown Medical Review

Vol. 8, Issue 1, 2024

Importance

Feedback is an essential part of graduate medical education. With increasing demands on time and fewer opportunities for formal feedback in residency, there is an increasing need for objective and actionable feedback.

Objective

To understand how the Crowd-Sourced Assessment of Technical Skills (CSATS) assessment of residents' laparoscopic skills compares with the current faculty assessment in an obstetrics/gynecology residency program.

Design, Setting, and Participants

This was a prospective educational intervention study involving residents from a single academic obstetrics/gynecology residency program at a large tertiary teaching hospital. Individual performance videos of hysterectomies, myomectomies, and endometriosis resections were recorded and uploaded to the CSATS database where they were analyzed. Formal written feedback reports from the CSATS system were sent to the residents. Residents then completed a survey addressing satisfaction, timeliness, and utility of the CSATS feedback method overall and in comparison with the residency's in-person feedback and with the standard written, electronic feedback evaluation.

Results

A total of 16 residents submitted 65 videos of operations. Overall comparison of CSATS vs faculty evaluation was equivalent. Residents found that CSATS response time was slower than receiving in-person feedback from faculty. When comparing CSATS with the standard feedback form, residents felt strongly that CSATS was superior for both surgical skill feedback and procedure evaluation, with both categories having a mean top score. When comparing low-volume CSATS utilization vs high-volume CSATS utilization, there were no statistical differences. Overall, residents reported high satisfaction with CSATS evaluation.

Conclusions

Residents were generally satisfied with CSATS as an overall feedback tool. CSATS could be a possible addition to help improve resident education.

Introduction

Feedback is fundamental to surgical training and has been found to be positively associated with performance improvement^{1,2} and ultimately patient outcomes after surgery.³ Residents have reported there are limited opportunities for feedback during training due to clinical pressures, duty hours, and decreased autonomy.^{4,5} As a result, there is an increasing need for a universal objective feedback system. Many training programs have started to implement universal rating tools to objectively assess skills.^{6,7} However, despite the development of good validated tools for objective surgical assessment, previous studies have shown that residents consider most of the feedback that they receive to be subjective, untimely, random, and nonactionable.^{8,9} This may be due to a lack of implementation of the validated tools.

Researchers in technical skill assessment in many fields have used video-recorded performances and expert reviewers to generate more data points for a single performance, lending more educational relevance to each individual performance.¹⁰⁻¹² Expert review is very time intensive and quite expensive, making routine expert review of video-recorded performance untenable for regular educational use. Crowd-Sourced Assessment of Technical Skills (CSATS) is a commercial program that uses the power of crowd-sourced internet reviewers to provide low-cost, rapid, data-rich assessment of surgical videos.¹⁰

If objective and comparable, CSATS may be a tool that will enable more impartial, actionable, quantitative assessment of trainees. We sought to understand how crowd-sourced assessment of residents' laparoscopic skills compares with current faculty assessment of residents' laparoscopic skills in a single obstetrics/gynecology (Ob/Gyn) residency program.

Materials and Methods

After obtaining approval from the MedStar Health institutional review board (No. CR00001361), senior residents (postgraduate years 3 and 4) from an academic Ob/Gyn program were recruited to participate in the study from September 2018 to January 2020. All residents rotating on the minimally invasive gynecological surgery rotation were invited to participate. Their performance of major laparoscopic gynecologic cases, including myomectomies, hysterectomies, and endometriosis excisions, were video recorded. To qualify for video submission, residents were required to complete more than 50% of the case. All cases were done with the programs' 2 minimally invasive gynecological surgery attendings. They could submit as many cases as they wanted. Individual performance videos were recorded and uploaded to the CSATS database, where they were analyzed, and written feedback reports were sent to the residents. During their cases, the residents received their standard intraoperative feedback from the attending physician

at the time of surgery, as well as the standard electronic feedback forms on the program's existing evaluation platform. The existing platform requires residents to send questionnaires to faculty after surgeries, asking for assessment. This standard feedback form is made up of 7 broad questions, only 1 of which titled "Surgical Technique," addresses surgical skill. The in-person feedback was not standardized. The attendings gave feedback as they deemed appropriate, both during the case and/or in a debrief after.

CSATS uses anonymous crowd-sourced internet reviewers to evaluate surgical performance. It assesses techniques using the validated Global Operative Assessment of Laparoscopic Skills (GOALS) that includes a 5-point Likert scale with verbal descriptors associated with 1, 3, and 5. Subjective written comments are also provided. Four of the GOALS domains (depth perception, bimanual dexterity, efficiency, and tissue handling) are separately assessed with the combination of all 4 averaged scores making up the final score (20 being the highest possible score). The fifth GOALS domain, autonomy, was not able to be assessed via video evaluation.

After completion of the video submission/feedback portion of the study, all residents were asked to complete a survey addressing satisfaction, timeliness, and utility of the CSATS feedback method overall and in comparison with the traditional attending feedback and the existing formal evaluation platform (see appendix for full questionnaire).

Results were analyzed by hospital system-employed statisticians. The mean and SD of the residents' review scores were computed overall. They were then broken into low usage of CSATS (1-4 videos) or high usage (>5 videos). The mean and SD for each group was computed and compared using a *t* test. The *P* values were calculated using the Wilcoxon signed-rank test. We considered a *P* value less than .05 to be significant.

Results

From September 2018 to August 2020, residents submitted videos to CSATS for evaluation. A total of 16 residents submitted 65 videos of operations. Of the 16 residents, 6 (37.5%) submitted 5 or more videos and 10 (62.5%) submitted 1 to 4 videos. All residents completed the feedback satisfaction survey once at the end of the study. Residents reported low completion of the existing electronic feedback form, with residents estimating less than 5% of their surgical cases received a written procedure evaluation (mean [SD], 4.3% [3.77%]).

All CSATS feedback was received by all residents within 1 week of submission. Sixty percent of feedback was received in less than 72 hours. The remaining 40% was received within 1 week ([Figure 1](#)).

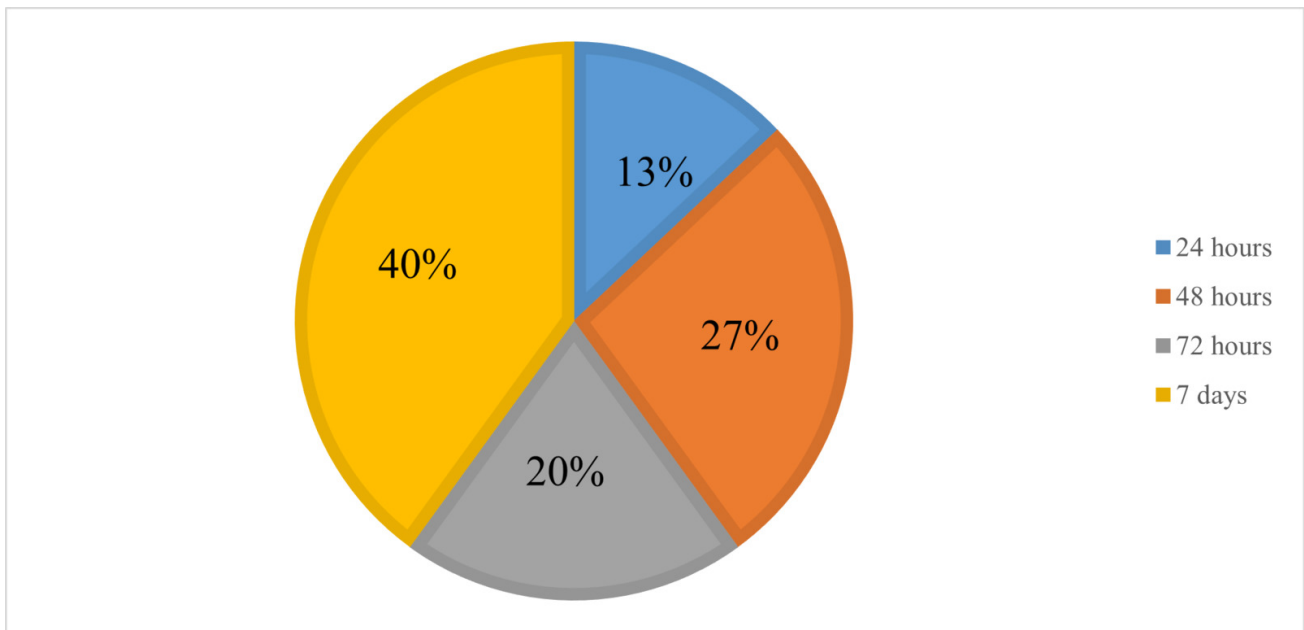


Figure 1. CSATS Average response time for 16 participants

The lowest scores on the survey were the questions comparing CSATS vs immediate, in-person faculty evaluation. Overall comparison of CSATS vs faculty evaluation scored a 3.25 (3 meaning CSATS and faculty evaluations were equivalent). Residents found that CSATS response time was slower than receiving in-person feedback from faculty (Table). We were unable to assess the response time in comparison with the standard evaluation form because of the low completion rate of the form.

When comparing CSATS with the standard feedback form, residents felt strongly that CSATS was superior to the standard electronic form for surgical skill feedback and procedure evaluation, with both categories receiving a mean top score. Residents agreed that CSATS was thorough, precise, and could help with laparoscopic skills (Table). However, the majority reported that they are not practicing laparoscopy more or less based on CSATS feedback.

When comparing responses from residents with low-volume (1-4 submitted videos) CSATS vs high-volume (≥ 5 videos) CSATS feedback, there were no statistical differences for any of the questions (Table).

Overall, residents reported satisfaction with CSATS evaluation, with a mean score of 4.13 out of 5. Residents reported that there may be an adjunctive role for CSATS as a standard tool for laparoscopic evaluation, with no statistical difference in low-volume compared with high-volume feedback ($P = .84$).

Table 1. Average Responses to the CSATS Questions

	Mean (SD)			<i>P</i> value ^A
	All	Low-volume CSATS Feedback	High-volume CSATS feedback	
CSATS alone				
How satisfied are you with CSATS evaluation of laparoscopic surgical skills?	4.13 (0.74)	4.11 (0.78)	4.17 (0.75)	.89
CSATS evaluation was thorough and precise	3.88 (0.72)	3.7 (0.67)	4.17 (0.75)	.22
CSATS evaluation will help improve my laparoscopic skills	3.62 (1.26)	3.5 (1.43)	3.83 (0.98)	.63
Are you exercising based on CSATS feedback and suggestions?	3 (0.97)	3 (0.94)	3 (1.10)	>.99
Should csats be implemented as a standard tool for laparoscopic evaluation in a residency program?	4.06 (0.93)	4.1 (0.88)	4 (1.10)	.84
CSats compared with in-person faculty evaluation				
CSATs evaluation compared with faculty evaluation overall	3.25 (1.48)	2.9 (1.60)	3.83 (1.17)	.24
CSATs evaluation compared with faculty evaluation time wise	2.81 (1.42)	3 (1.63)	2.5 (1.05)	.52
Csats compared with standard feedback form				
CSATS compared with the standard feedback form procedure evaluation	4.67 (0.62)	4.78 (0.44)	4.5 (0.84)	.41
How do you compare surgical skills feedback you received with csats vs the standard feedback form overall	4.56 (0.73)	4.6 (0.7)	4.5 (0.84)	.80

Abbreviation: CSATS, Crowd- Sourced Assessment of Technical Skills.

^A*P* value compares low-volume vs high-volume responses.

Discussion

As changes in residency training have evolved, we have to continue to create new education strategies to help maximize trainees' learning. Decreasing work hours and fewer opportunities^{4,5} in the operating department mean we must maximize our opportunities for effective feedback. Feedback has been shown to improve efficiency, safety, and skill.¹ However, at our training hospital, laparoscopic education and evaluation have limitations. Feedback is largely subjective and varies from intraoperative feedback to general feedback at the end of a 5-week rotation. Residents have to request formal evaluations after individual cases. This process is not a residency requirement, and it is therefore infrequently completed. This likely contributed to the residents' increased satisfaction with CSATS over the standard written feedback. The goal of the study was to investigate how crowd-sourced assessment of residents' laparoscopic skills compares with faculty assessment of residents' laparoscopic skills in an Ob/Gyn residency program.

Previous studies have shown that CSATS is an accurate, acceptable method for objective technical skills assessment.¹¹ Residents were generally satisfied with CSATS as a form of feedback, both in timeliness and as an overall feedback tool. In terms of timeliness, all results were back within 7 days. This is important because previous studies have found that feedback given more than 7 days out is perceived as less effective.¹³ While residents agreed that CSATS should be implemented as a standard tool for evaluation, most found it was comparable with, but not better than, traditional real-time and in-person feedback. They strongly felt that it was superior to the standard feedback form. This is consistent with previous studies that have shown residents find written feedback to be a good adjunct to face-to-face feedback, but find face-to-face feedback more important to their education.¹³

Residents felt strongly that CSATS as a form of feedback was far better than the existing written system. The top scores that CSATS received in comparison with the existing form demonstrates the value that CSATS can add to resident surgical education in terms of formal feedback. Formal, reviewable feedback has been found to be fundamental for resident assessment and is an important component of resident education.¹⁴ A more longitudinal study of objective resident improvement after the use of CSATS could help solidify its value as a feedback tool for residents.

Despite their satisfaction with CSATS, residents did not feel they would implement new practices or exercises based on their CSATS feedback. This suggests that the practical implications of CSATS may be limited. However, in this study, residents only used CSATS for a short period, which may be why they did not consistently implement the feedback. A more longitudinal study in which residents use it consistently may show better utilization of the feedback.

Additional studies in the future may include development of new validated simulation exercises that predictably drive improvement in operative skill deficits as assessed by CSATS as well as following skill improvement by tracking CSATS scores. The long-term goal of studying CSATS in the educational arena is to drive improvement in surgical assessment, skills acquisition, patient safety, and quality surgical outcome measures. This research also demonstrated the importance to resident education of timely in-person feedback from clinical faculty.

A significant disadvantage of CSATS is the associated cost. The cost of implementing CSATS for 1 year is approximately \$30 000. This can be a burden to the graduate medical education budget. Educational materials make up approximately one-third of the hidden costs of graduate medical education that are not reimbursed by the federal government.¹⁵ This could be a significant hurdle that many programs would have difficulty overcoming, especially given our findings of equivalent satisfaction with standard face-

to-face feedback. However, if future studies explored resident surgical skill acquisition after using CSATS and saw an improvement in those who utilized it, the cost could be justified.

There were limitations to this study. This was a small study with only 16 participants, which increases the chance of type 2 error. Implementation of CSATS to the program was difficult. Almost two-thirds of the participants submitted fewer than 5 videos, indicating poor participation. However, there was not a significantly higher score in any category for those residents who submitted 5 or more videos, implying that there may not be increased satisfaction or value for those who use it more frequently.

While the study provided insight into residents' opinion regarding CSATS compared with the program's standard feedback form, the 2 feedback methods used different criteria for evaluation. CSATS uses the GOALS criteria, while the existing feedback form merely had 1 question referencing surgical technique. It is possible that the preference towards CSATS may actually be a preference towards the GOALS criteria. We did not change the existing feedback form to use the GOALS criteria for this study because we wanted a comparison of CSATS with what was already in use. A future study comparing GOALS criteria in a standard feedback form vs CSATS could potentially provide a more direct comparison of the 2 different evaluations. This study was additionally limited by the low response rate of the existing feedback form. Therefore, there may have been a bias response when comparing against CSATS, where trainees will always have a 100% response rate. However, this limitation could also be considered a strength for a tool like CSATS that allows for consistent and reliable feedback.

To our knowledge, there are no previous studies that have evaluated CSATS utility in an Ob/Gyn residency program or that assessed resident satisfaction with CSATS as a tool.

Conclusions

While the standard method of face-to-face feedback is invaluable, CSATS may be of value as an adjunct tool for trainee feedback, and it may improve on some existing written and electronic surgical evaluation systems.

Acknowledgments

We thank the statisticians at MedStar Health Research Institute for their statistical support.

Disclosures

The authors report no conflicts of interest.

Sources of Funding

This study was funded by MedStar Washington Hospital Center Graduate Medical Education.

Meeting Presentation

This study was presented at the 49th Annual Scientific Meeting of Society of Gynecologic Surgeons; March 20th, 2023; Tucson, Arizona.

Submitted: June 17, 2024 EST. Accepted: August 20, 2024 EST. Published: October 28, 2024 EST.

REFERENCES

1. Trehan A, Barnett-Vanes A, Carty MJ, McCulloch P, Maruthappu. The impact of feedback of intraoperative technical performance in surgery: a systematic review. *BMJ Open*. 2015;5:6. doi:[10.1136/bmjopen-2014-006759](https://doi.org/10.1136/bmjopen-2014-006759)
2. Oestergaard J, Bjerrum F, Maagaard M, et al. Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized educational trial. *BMC Med Educ*. 2012;12:7. doi:[10.1186/1472-6920-12-7](https://doi.org/10.1186/1472-6920-12-7)
3. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med*. 2013;369:1434-1442. doi:[10.1056/NEJMsa1300625](https://doi.org/10.1056/NEJMsa1300625)
4. Ong CC, Dodds A, Nestel D. Beliefs and values about intra-operative teaching and learning: a case study of surgical teachers and trainees. *Adv Health Sci Educ Theory Pract*. 2016;21(3):587-607. doi:[10.1007/s10459-015-9654-5](https://doi.org/10.1007/s10459-015-9654-5)
5. Hashimoto DA, Bynum WE, Lillemoe KD, Sachdeva AK. See more, do more, teach more: surgical resident autonomy and the transition to independent practice. *Acad Med*. 2016;91(6):757-760. doi:[10.1097/ACM.0000000000001142](https://doi.org/10.1097/ACM.0000000000001142)
6. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Eng J Med*. 2006;355(25):2664-2669. doi:[10.1056/NEJMra054785](https://doi.org/10.1056/NEJMra054785)
7. Kramp KH, Van Det MJ, Hoff C, et al. Validity and reliability of Global Operative Assessment of Laparoscopic Skills (GOALS) in novice trainees performing laparoscopic cholecystectomy. *J Surg Educ*. 2015;72(2):351-358. doi:[10.1016/j.jsurg.2014.08.006](https://doi.org/10.1016/j.jsurg.2014.08.006)
8. Levinson KL, Barlin JN, Altman K, Stain AJ. Disparity between resident and attending physician perceptions of intraoperative supervision and education. *J Surg Educ*. 2010;2(1):31-36. doi:[10.4300/JGME-D-09-00096.1](https://doi.org/10.4300/JGME-D-09-00096.1)
9. Rose JS, Waibel BH, Schenarts PJ. Disparity between resident and faculty surgeons' perceptions of preoperative preparation, intraoperative teaching, and post operative feedback. *J Surg Educ*. 2011;68(6):459-464. doi:[10.1016/j.jsurg.2011.04.003](https://doi.org/10.1016/j.jsurg.2011.04.003)
10. Chen C, White L, Kowalewski T, et al. Crown-Sourced Assessment of Technical Skills: a novel method to evaluate surgical performance. *J Surg Res*. 2015;71(2):351-358.
11. Holst D, Kowalewski T, White L. Crown-Sourced Assessment of Technical Skills: an adjunct to urology resident surgical simulation training. *J Endourol*. 2015;29(5):604-609. doi:[10.1089/end.2014.0616](https://doi.org/10.1089/end.2014.0616)
12. Vernez SL, Huynh V, Osann K, Okjunov Z, Landman J, Clayman RV. C-SATS: assessing surgical skills among urology residency applicants. *J Endourol*. 2017;31(S1):S95-S100. doi:[10.1089/end.2016.0569](https://doi.org/10.1089/end.2016.0569)
13. Bello RJ, Sarmiento S, Meyer ML, et al. Understanding surgical resident and fellow perspectives on their operative performance feedback needs: a qualitative study. *J Surg Educ*. 2018;75(6):1498-1503. doi:[10.1016/j.jsurg.2018.04.002](https://doi.org/10.1016/j.jsurg.2018.04.002)
14. Hartranft TH, Yandle K, Graham T, Holden C, Chambers LW. Evaluating surgical residents quickly and easily against the milestones using electronic formative feedback. *J Surg Educ*. 2017;74(2):237-242. doi:[10.1016/j.jsurg.2016.09.006](https://doi.org/10.1016/j.jsurg.2016.09.006)
15. Kelly SP, Tibbles C, Barnett SR, Schwartzstein RM. The “hidden costs” of graduate medical education in the United States. *J Grad Med Educ*. 2012;4(2):267-268. doi:[10.4300/JGME-D-12-00038.1](https://doi.org/10.4300/JGME-D-12-00038.1)

SUPPLEMENTARY MATERIALS

Appendix

Download: <https://gmr.scholasticahq.com/article/124417-crowd-sourced-assessment-of-laparoscopic-surgical-skills-for-ob-gyn-residents/attachment/248259.docx>
