



Global assessment of surgical skills (GASS): validation of a new instrument to measure global technical safety in surgical procedures

Peter Nau¹ · Erin Worden¹ · Ryan Lehmann¹ · Kyle Kleppe² · Gregory J. Mancini² · Matt L. Mancini² · Bruce Ramshaw^{3,4}

Received: 30 March 2023 / Accepted: 8 May 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Background Broad implementation of the American Board of Surgery’s entrustable professional activities initiative will require assessment instruments that are reliable and easy to use. Existing assessment instruments of general laparoscopic surgical skills have limited reliability, efficiency, and validity across the spectrum of formative (low-stakes) and summative (high-stakes) assessments. A novel six-item global assessment of surgical skills (GASS) instrument was developed and evaluated with a focus upon safe versus unsafe surgical practice scoring rubric.

Methods The GASS was developed by iterative engagement with expert laparoscopic surgeons and includes six items (economy of motion, tissue handling, appreciating operative anatomy, bimanual dexterity, achievement of hemostasis, overall performance) with a uniform three-point scoring rubric (“poor–unsafe”, “adequate–safe”, “good–safe”). To test inter-rater reliability, a cross-sectional study of four bariatric surgeons with experience ranging from 4 to 28 years applied the GASS and the global operative assessment of laparoscopic skills (GOALS) to 30 consecutive Roux-en-Y gastric bypass procedure operative videos. Inter-rater reliability was assessed for a simplified dichotomous “safe” versus “unsafe” scoring rubric using Gwet’s AC₂.

Results The GASS inter-rater reliability was very high across all six domains (0.88–1.00). The GASS performed comparably to the GOALS inter-rater reliability scores (0.96–1.00). The economy of motion and bimanual dexterity items had the highest percentage of unsafe ratings (9.2% and 5.8%, respectively).

Conclusion The GASS, a novel six-item instrument of general laparoscopic surgical skills, was designed with a simple scoring rubric (poor-safe, adequate-safe, good-safe) to minimize rater burden and focus feedback to trainees and promotion evaluations on safe surgical performance. Initial evaluation of the GASS is promising, demonstrating high inter-rater reliability. Future research will seek to assess the GASS against a broader spectrum of laparoscopic procedures.

Keywords Clinical competence · General surgery · Surgical procedures, operative · Video tape recording · Gastric bypass · Process assessment, health care

While operative techniques are only one component of overall surgical competency, it is one of the most challenging

to measure [1]. Traditional assessment methods of operative logs, end-of-rotation global assessments, attestations of program directors, and direct observation without objective criteria are unreliable [1, 2] and are subject to concerns regarding gender bias and inequity within surgical training programs [3, 4].

The adoption of competency-based medical education (CBME) and entrustable professional activities (EPAs) by the America Board of Surgery deliberately moves from time-based to competency-based requirements to ensure graduating surgical residents are uniformly prepared for autonomous practice. The core-operating components of this educational reform include the use of micro-assessments

✉ Peter Nau
peter-nau@uiowa.edu

¹ Department of Surgery, Section of Bariatric Surgery, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, Iowa City, IA 52242, USA

² Department of Surgery, Section of Foregut Surgery, University of Tennessee, Knoxville, TN, USA

³ CQInsights PBC, Knoxville, TN, USA

⁴ Caresyntax Corporation, Boston, MA, USA

used in routine surgical practice, the provision of regular meaningful feedback, and summative documentation to inform trainee promotion decisions [2]. Documented evaluation with CBME and EPAs outcomes is limited but early experience [5, 6] has demonstrated increased burden upon mentors (administrative) and trainees (psychological) [7, 8]. Successful implementation of CBME and EPAs will rely heavily on reliable, valid, bias-free assessment instruments that can be used efficiently within the context of a busy surgical practice [9].

The growth of minimal invasive surgery since the 1980's has been steady and in 2018, 18% of ambulatory surgical procedures were performed via minimally invasive techniques (i.e., robotic or laparoscopic) [10, 11]. Achieving mastery for minimally invasive surgeries can be characterized by long learning curves (30–80 cases) which may require years to achieve [12–14]. To achieve CBME and EPAs in minimally invasive surgery, a variety of competency assessment instruments have been developed including the laparoscopic suturing competency assessment tool (LS-CAT), global rating index of technical skills (GRITS), global operative assessment of laparoscopic skills (GOALS), global evaluative assessment of robotic skills (GEARS), and objective structured assessment of technical skills (OSATS) [15–21]. These instruments have several important features in common including (a) designed to assess level of competency of surgical technique, (b) scoring rubrics characterized by descriptive responses that routinely combine different aspects of surgical performance (non-mutually exclusive) that can complicate scoring decisions, and (c) minimal evidence supporting their use for high-stakes summative assessments [22–26]. The GOALS and OSATS have had reported difficulties with inter-rater reliability [22, 23] and the GEARS and LC-CAT have relatively limited focus for minimal invasive surgery assessments (i.e., robotic and suturing skill). The GOALS assesses competency with depth perception, bimanual dexterity, efficiency, and tissue handling and includes an assessment of trainee autonomy during the procedure and level of case difficulty [17]. Each of the GOALS items is scored on a five-point scale with unique anchor descriptors describing the level of proficiency with technique.

Perhaps most importantly, none of the above instruments made use of a framework to explicitly assess surgical skills as safe or unsafe, a fundamentally important goal of surgical training. The objective of this research was to provide preliminary evidence of the face validity and test–retest reliability of the scoring rubric of a novel six-item general skills assessment focused on safe versus unsafe surgical technique.

Methods

This study was a one-time assessment of 30 consecutive, de-identified, and randomized Roux-en-Y Gastric Bypass (RYGB) procedure operative videos, scored by four expert bariatric surgeons using the GASS instrument.

Instrument development

Four expert surgeons (PN, EW, RL, BR) developed the instrument using a real-world evidence approach including rapid iterative design review and revision, regular testing of instrument constructs against real-world surgical videos, and consistent appraisal of the instrument's ability to achieve its intended use (i.e., assessment of safe surgical practice). To address the limitations of existing instruments, this surgical author team designed the instrument to assess six domains of intraoperative surgical practice including (1) economy of motion, (2) tissue handling, (3) appreciating operative anatomy, (4) bimanual dexterity, (5) achievement of hemostasis, and (6) overall performance. A priori instrument design considerations included (a) focus upon safe surgical performance, (b) unambiguous scoring criteria including mutually exclusive assessment categories, and (c) amenable to minimally invasive intraoperative video-based assessments. Each of the six items was scored on a three-point scale: (a) poor–unsafe, (b) adequate–safe, and (c) good–safe (Fig. 1).

Data collection

A sample of 30 consecutive laparoscopic RYGB surgical videos conducted between August 2021 and January 2022 was used for the assessments. Videos were uploaded to a

Domain	Response Options
Economy of motion	1. Poor (unsafe) 2. Adequate (safe) 3. Good (safe)
Tissue handling	1. Poor (unsafe) 2. Adequate (safe) 3. Good (safe)
Appreciating operative anatomy	1. Poor (unsafe) 2. Adequate (safe) 3. Good (safe)
Bimanual dexterity	1. Poor (unsafe) 2. Adequate (safe) 3. Good (safe)
Achievement of hemostasis	1. Poor (unsafe) 2. Adequate (safe) 3. Good (safe)
Overall Performance	1. Poor (unsafe) 2. Adequate (safe) 3. Good (safe)

Fig. 1 The GASS instrument

proprietary, online system to which four bariatric surgeon reviewers ranging in experience from 4 to 28 years of practice serving as faculty for minimal invasive surgery-bariatric surgery training programs rated the jejunojejunostomy (JJ) portion of the RYGB procedure. Reviewers were blinded to other reviewer assessments. Raters used a pre-formatted Excel spreadsheet to rate each of the 30 videos using the GASS instrument. For comparison, raters also completed the GOALS assessment and a single-item measure of global case difficulty (easy, medium, hard) for the 30 videos. Raters were provided scoring sheets and received no specific training in the completion of assessments. Operative videos were randomized for presentation to the reviewers.

Analysis

Inter-rater reliability was assessed using the 2-point safe–unsafe scoring rubric, in which the responses of adequate and good were collapsed into the safe category. Gwet's AC_2 and total agreement were used to assess inter-rater reliability. Total agreement is the ratio of raters that agreed on case safety as either safe or unsafe and ranges from 50 to 100%. Gwet's AC_2 is a weighted measure that adjusts for scales that have a high probability of random agreement and that allows for partial agreement between categories. Gwet's AC_2 was calculated using quadratic weights, with a score range from 0 to 100%. In the Supplemental Tables, results are presented for the full 3-point GASS scale including Cronbach's alpha, which was used as measure of internal consistency across the four reviewers and for Spearman's rank correlation coefficient between each individual item of the GOALS, GASS, and global case difficulty score.

Results

All four raters completed the six-item GASS for the 30 surgical videos (720 individual item assessments). As measured by Gwet's AC_2 , the inter-rater reliability was very high for all six of the GASS items (0.88–1.00) (Table 1). In general,

Table 1 Inter-rater reliability of the GASS instrument

	Gwet's AC_2	Total agreement (%)
Overall performance	0.947	97.5
Economy of motion	0.820	90.8
Achievement of hemostasis	1.000	100.0
Appreciating operative anatomy	0.947	97.5
Bimanual dexterity	0.881	94.2
Tissue handling	0.929	96.7

the GASS inter-rater reliability remained stable for different levels of case difficulty. The inter-rater reliability for appreciating operative anatomy, bimanual dexterity, tissue handling, and overall performance had minor differences in Gwet's AC_2 between cases with higher versus lower levels of case difficulty. One item (economy of motion) showed greater inter-rater reliability for higher case difficulties. The GASS inter-rater reliability performed comparably to the GOALS inter-rater reliability which was also very high for all four GOALS items as measured by Gwet's AC_2 (0.96–1.00). The GASS Cronbach's alpha ranged from 0.71 to 0.76 suggesting high internal consistency (Supplemental Table 1).

Among the four reviewers, the average rating using the 3-point scale, was highest for achievement of hemostasis (2.8), followed by appreciating operative anatomy (2.5), tissue handling (2.4), bimanual dexterity (2.3), and economy of motion (2.2) (Fig. 1). The item with the highest percent of unsafe ratings was economy of motion (9.2%), followed by bimanual dexterity (5.8%), and then tissue handling (3.3%). The achievement of hemostasis had no unsafe ratings (Fig. 2).

The correlation between GASS and GOALS items was modest, ranging between 0.12 and 0.79 (Supplemental Table 2). The lowest correlation was noted from achievement of hemostasis, which consistently received high scores with limited variability. Among the remaining items, the lowest correlation was 0.34 (economy of motion correlated with tissue handling). Of particular note, the two similar items between the GASS and GOALS, specifically bimanual dexterity and tissue handling, did not have perfect (or near

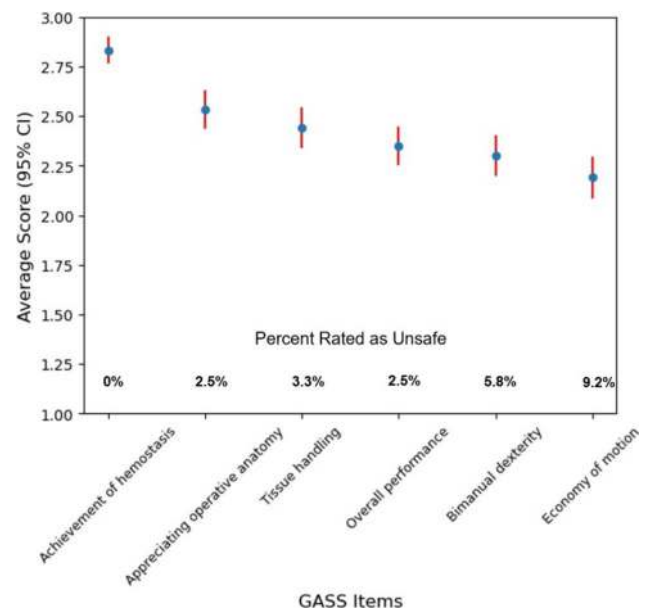


Fig. 2 Mean and 95% confidence interval for the individual items in the GASS

perfect) correlations indicating that the scoring rubric was a significant contributor to differences in scoring. The correlation between the two tissue handling items was 0.77 and between the two bimanual dexterity items was 0.79 (Supplemental Table 2). Both the GOALS and the GASS total scores were modestly inversely associated with global case difficulty, indicating that surgical performance scores were lower among more difficult cases (Table 2).

Discussion

Current standardized instruments to assess general surgical technical performance (e.g., OSATS, GOALS, GEARS, among others) are scored based upon perceived *level of technical skill and knowledge* in domains such as instrument handling, efficiency of motion, flow of operation, and respect for tissue measured in terms of subjective assessment. Reorienting assessments of surgical quality to focus upon perceived *level of safety* has the potential to serve as a more objective basis for credentialing and improvement activities.

The six-item GASS instrument evaluated in this study was designed with that objective in mind, to focus on safe surgical practice. The preliminary results reported in this study indicate that the inter-rater reliability of scoring was in the very good to excellent range with stability across case difficulty types with results that are comparable to or better than other validated instruments designed to measure surgical skill [17, 20].

Of the six items, the achievement of hemostasis resulted in a consistently high score, with limited variability. This was further reflected in a low correlation with other response options and no “unsafe” ratings, as no surgeon would close an elective surgical patient without achieving hemostasis. Upon review of these results and engagement among the surgeon authors, this item may better be scored as a simple “yes” or “no” response option. In other words, hemostasis is

either achieved or it is not. As the latter is an outcome that should almost never happen and because of the importance of identifying outlier performance on this critical issue, the consensus was to retain this item in the next iteration of the GASS instrument but not rate it as safe or unsafe but rather as achieved or not achieved. Alternatively, the hemostasis item could be modified to assess “Skill in achieving hemostasis,” which would enable a rater to score a surgeon as “poor” based upon unreasonable or excessive blood loss that should and could have been controlled or prevented; “adequate” if blood loss was acceptable and controlled and could be improved, and “expert” if blood loss was minimal.

Existing surgical skill assessment instruments have failed to achieve widespread adoption in real-world training programs. The successful implementation of CBME and EPAs requires instruments that (a) have clinical face validity, (b) can be reliably scored, (c) can be used efficiently within the context of busy surgical training programs or practice, (d) provide results that translate into specific, actionable guidance to trainees, and (e) lead to the graduation of competent surgeons that are fully prepared for autonomous practice. The GASS instrument was designed specifically with a simplified scoring rubric (safe vs. unsafe) that can be applied efficiently in busy environments for formative (low-stake) assessments designed to improve trainee performance as well as summative (high-stakes) assessments designed to ensure safe promotion to autonomous practice. The GASS was able to achieve high levels of inter-rater reliability with no training provided to raters, a positive result that suggests the safe–unsafe framework has the potential to scale to other types of procedures.

Future research into the utility and validity of the GASS should include a qualitative assessment of its utility by surgeon trainers and of the value of the instrument to the trainee and expanded investigation into the instrument’s reliability and utility in other surgical contexts (other procedures, other training programs, etc.). Importantly, additional research to understand the perspective of expert surgeon raters on

Table 2 Inter-rater reliability performance of the GASS instrument by average global case difficulty score

	Less difficult case ^a		More difficult case ^b	
	Gwet’s AC ₂	Total agreement (%)	Gwet’s AC ₂	Total agreement (%)
Overall performance	1.00	100.0	0.904	95.6
Economy of motion	0.716	86.5	0.890	94.1
Achievement of hemostasis	1.000	100.0	1.000	100.0
Appreciating operative anatomy	1.000	100.0	0.904	95.6
Bimanual dexterity	0.917	96.2	0.852	92.6
Tissue handling	0.960	98.1	0.904	95.6

^aAverage global case difficulty score < 2

^bAverage global case difficulty score ≥ 2

the specific issues that warrant assigning an unsafe rating will contribute to an improved instruction set and training materials.

The American Board of Surgery's EPA initiative is an important advance in trainee evaluation and a challenging implementation that requires investment by surgical training programs to optimize impact [9, 26, 27]. Video-based assessments (VBAs) are increasingly being used for formative and summative assessments and will likely continue to play an important and evolving role in promotion from laparoscopic-based surgical residencies, fellowships, and board certification [28, 29]. The GASS provides a straight-forward instrument that can be used in combination with VBAs to support EPA implementation.

Strengths

The six-item GASS instrument was designed by a team of four surgeons with expertise in bariatric surgery. The preliminary scoring, and subsequent reliability testing, of the GASS instrument were conducted by four expert reviewers with a diverse range of experience from two large university systems. The scoring of the GASS was conducted using a methodology that is consistent with the intended use of the tool, specifically as part of VBAs of operative videos from real-world surgical practice. Finally, simultaneous reviews of the GOALS and GASS enabled performance comparison against existing established tools with comparable inter-rater reliability.

Limitations

Perhaps the most important limitation of the study was the selection of videos and the relatively few unsafe ratings. The process of video selection, while acceptable from a research perspective, (30 consecutive procedures), may have resulted in a limited spectrum of operative case difficulty and variation. Only one type of surgical procedure was used for the study which also limits comment regarding the generalizability to other procedures. Future research efforts will seek to assess the GASS performance across multiple types of surgical procedures as well as a greater spectrum of surgical case difficulty.

Conclusion

Creating objective, reproducible, easy to implement assessments of surgical safety is a critical enabler of CBME and EPA. The GASS is a new, reliable instrument that assesses intraoperative safety of general surgical skills and has the potential to support surgical training and professional credentialing.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00464-023-10116-8>.

Acknowledgements The authors wish to thank Justin Smith for analytical support and Health Analytics, LLC for administrative and medical writing support in preparing the manuscript.

Funding Caresyntax was the sponsor for this project.

Declarations

Disclosures Dr. Nau reported receipt of support for this manuscript and payment or honoraria from Caresyntax. Dr. Warden reported receipt of support for this manuscript from Caresyntax. Dr. Lehmann reported receipt of support for this manuscript, consulting fees, and payment or honoraria from Caresyntax. Dr. Kleppe reported receipt of support for this manuscript from Caresyntax. Dr. G Mancini reported receipt of support for this manuscript, consulting fees, and payment or honoraria from Caresyntax. Dr. M Mancini reported receipt of support for this manuscript, consulting fees, and payment or honoraria from Caresyntax. Dr. Ramshaw reported receipt of support for this manuscript and consulting fees from Caresyntax.

References

1. Tsue TT, Dugan JW, Burkey B (2007) Assessment of surgical competency. *Otolaryngol Clin North Am.* <https://doi.org/10.1016/j.otc.2007.07.005>
2. American Board of Surgery (2022) Entrustable professional activities. Available from <https://www.absurgery.org/default.jsp?epahome>. Accessed 1 Mar 2023
3. Gerull KM, Loe M, Seiler K, McAllister J, Salles A (2019) Assessing gender bias in qualitative evaluations of surgical residents. *Am J Surg.* <https://doi.org/10.1016/j.amjsurg.2018.09.029>
4. Dill-Mackay A, Hsu CH, Neumayer LA, Nfonam VN, Turner AP (2022) The role of implicit bias in surgical resident evaluations. *J Surg Educ.* <https://doi.org/10.1016/j.jsurg.2021.12.003>
5. Bhatti NI, Cummings CW (2007) Competency in surgical residency training: defining and raising the bar. *Acad Med.* <https://doi.org/10.1097/ACM.0b013e3180555bfb>
6. Gauvin G, Hay K, Hopman W, Hurton S, Lim S, Zevin B, Jalink D, Nanji S (2021) Competency-based education in general surgery: are Canadian residents ready? *Can J Surg.* <https://doi.org/10.1503/cjs.011520>
7. Cheung K, Rogoza C, Chung AD, Kwan BYM (2022) Analyzing the administrative burden of competency based medical education. *Can Assoc Radiol J.* <https://doi.org/10.1177/08465371211038963>
8. Ott MC, Pack R, Cristancho S, Chin M, Van Koughnett JA, Ott M (2022) "The most crushing thing": understanding resident assessment burden in a competency-based curriculum. *J Grad Med Educ.* <https://doi.org/10.4300/JGME-D-22-00050.1>
9. Champagne BJ (2013) Effective teaching and feedback strategies in the OR and beyond. *Clin Colon Rectal Surg.* <https://doi.org/10.1055/s-0033-1356725>
10. St John A, Caturegli I, Kubicki NS, Kavic SM (2020) The rise of minimally invasive surgery: 16 year analysis of the progressive replacement of open surgery with laparoscopy. *JLS.* <https://doi.org/10.4293/JLS.2020.00076>
11. Mattingly AS, Chen MM, Divi V, Holsinger FC, Saraswathula A (2023) Minimally invasive surgery in the United States, 2022: understanding its value using new datasets. *J Surg Res.* <https://doi.org/10.1016/j.jss.2022.08.006>

12. Harrysson IJ, Cook J, Sirimanna P, Feldman LS, Darzi A, Aggarwal R (2014) Systematic review of learning curves for minimally invasive abdominal surgery: a review of the methodology of data collection, depiction of outcomes, and statistical analysis. *Ann Surg*. <https://doi.org/10.1097/SLA.0000000000000596>
13. Van Workum F, Stenstra MHB, Berklemans GHK, Slamán AE, van Berge Henegouwen MI, Gisbertz SS, van den Wildenberg FJH, Polat F, Irino T, Nilsson M, Nieuwenhuijzen GAP, Luyer MD, Adang EM, Hannink G, Rovers MM, Rosman C (2019) Learning curve and associated morbidity of minimally invasive esophagectomy: a retrospective multicenter study. *Ann Surg*. <https://doi.org/10.1097/sla.0000000000002469>
14. Weksler B (2018) Defining the learning curve for minimally invasive segmentectomy. *J Thorac Cardiovasc Surg*. <https://doi.org/10.1016/j.jtcvs.2018.06.065>
15. Doyle DD, Webber EM, Sidhu RS (2007) A universal global rating scale for the evaluation of technical skills in the operating room. *Am J Surg*. <https://doi.org/10.1016/j.amjsurg.2007.02.003>
16. Ijgosse WM, Leijte E, Ganni S, Luursema JM, Francis NK, Jakimowicz JJ, Botden SMBI (2020) Competency assessment tool for laparoscopic suturing: development and reliability evaluation. *Surg Endosc*. <https://doi.org/10.1007/s00464-019-07077-2>
17. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D, Fried GM (2005) A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg*. <https://doi.org/10.1016/j.amjsurg.2005.04.004>
18. Ferriss JS, Frost AS, Heinzman AB, Tsai R, Patterson D, Patzkowsky K, Blanck J, Bienstock JL (2021) Systematic review of intraoperative assessment tools in minimally invasive gynecologic surgery. *J Minim Invasive Gynecol*. <https://doi.org/10.1016/j.jmig.2020.10.007>
19. Hatala R, Cook DA, Brydges R, Hawkins R (2015) Constructing a validity argument for the objective structured assessment of technical skills (OSATS): a systematic review of validity evidence. *Adv Health Sci Educ Theory Pract*. <https://doi.org/10.1007/s10459-015-9593-1>
20. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. <https://doi.org/10.1046/j.1365-2168.1997.02502.x>
21. Goh AC, Goldfarb DW, Sander JC, Miles BJ, Dunkin BJ (2012) Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. *J Urol*. <https://doi.org/10.1016/j.juro.2011.09.032>
22. Kramp KH, van Det MJ, Hoff C, Lamme B, Veeger NJGM, Pierie JPEN (2015) Validity and reliability of global operative assessment of laparoscopic skills (GOALS) in novice trainees performing a laparoscopic cholecystectomy. *J Surg Educ*. <https://doi.org/10.1016/j.jsurg.2014.08.006>
23. Gallagher AG, O'Sullivan GC, Leonard G, Bunting BP, McGlade KJ (2014) Objective structured assessment of technical skills and checklist scales reliability compared for high stakes assessments. *ANZ J Surg*. <https://doi.org/10.1111/j.1445-2197.2012.06236.x>
24. Jarocki A, Rice D, Kent M, Oh D, Lin J, Reddy RM (2022) Validity of robotic simulation for high-stakes examination: a pilot study. *J Robot Surg*. <https://doi.org/10.1007/s11701-021-01258-9>
25. Goldenberg MG, Lee JY, Kwong JCC, Grantcharov TP, Costello A (2018) Implementing assessments of robot-assisted technical skill in urological education: a systematic review and synthesis of the validity evidence. *BJU Int*. <https://doi.org/10.1111/bju.14219>
26. Ryan JF, Mador B, Lai K, Campbell S, Hyakutake M (2022) Validity evidence for procedure-specific competence assessment tools in general surgery: a scoping review. *Ann Surg*. <https://doi.org/10.1097/SLA.0000000000005207>
27. Mayne A, Wilson L, Kennedy N (2020) The usefulness of procedure-based assessments in postgraduate surgical training within the intercollegiate surgical curriculum programme; a scoping review. *J Surg Educ*. <https://doi.org/10.1016/j.jsurg.2020.03.005>
28. McQueen S, McKinnon V, VanderBeek L, McCarthy C, Sonnadara R (2019) Video-based assessment in surgical education: a scoping review. *J Surg Educ*. <https://doi.org/10.1016/j.jsurg.2019.05.013>
29. Huang RJ, Limsui D, Triadafilopoulos G (2018) Video-based performance assessment in endoscopy: moving beyond “see one, do one, teach one”? *Gastrointest Endosc*. <https://doi.org/10.1016/j.gie.2017.09.014>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.