Video-Based Assessment in Surgical Education: A Scoping Review

Sydney McQueen, BSc, MSc,* Victoria McKinnon, BASc, MSt,† Laura VanderBeek, MD,† Colm McCarthy, MDCM, MSc,† and Ranil Sonnadara, PhD*,†

*Department of Surgery, University of Toronto, Toronto, Ontario, Canada; and †Department of Surgery, McMaster University, Hamilton, Ontario, Canada

BACKGROUND AND OBJECTIVE: Video-based assessment of residents' surgical skills may offer several advantages over direct observations of clinical performance in terms of objectivity, time-efficiency, and feasibility. Although video-based assessment is becoming more common in surgical training, a broad understanding of its utility is lacking. This scoping review explores video-based assessment in surgical training and presents the evidence supporting its use.

DESIGN: A literature search was conducted using the Web of Science database with key words related to video-based assessment and surgical training. Exclusion criteria included articles not published in English and articles on undergraduate medical education, continuing professional development, or non-surgical disciplines. Initially, 702 articles were identified; after title, abstract, and full-text screening by two independent reviewers (SM and VM), 199 articles remained.

RESULTS: We present the benefits of video-based assessment, including the ability to capture clinical ability in the operating room without decreasing intraoperative efficiency, as well as the potential to improve formative assessment and feedback practices. We describe the validity, reliability, and challenges of video-based assessment, as well as the use of video-based methods in clinical and simulated settings. We conclude by discussing questions that remain to be addressed.

CONCLUSIONS: Although further research and cost-benefit analyses are required, greater adoption of video-based assessment into surgical training may help meet increased assessment demands in an era of competency-based medical education. (J Surg Ed 76:1645−1654. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: Video-based assessment, surgical education, feedback, competency-based medical education

COMPETENCIES: Practice-Based Learning and Improvement, Patient Care, Medical Knowledge, Interpersonal and Communication Skills, Professionalism

INTRODUCTION

The shift to competency-based medical education (CBME) in Canada, the United States, and across the globe necessitates more robust assessment practices in order to determine when trainees have become proficient in specific required skills and to guide the development of competence through feedback.[1,2] In particular, formative assessment must become more fully and meaningfully integrated into all aspects of training and used to enhance the learning process through timely and constructive feedback to trainees.[3,4]

Educators have begun responding to these needs by developing a myriad of new assessment tools,[5] implementing portfolio-based approaches,[6] and moving toward programmatic models of assessment which require more frequent evaluations of trainees.[7,8] Advances have also come out of the United Kingdom with the development of the Orthopedic Competence Assessment Project[9] and the Intercollegiate Surgical Curriculum Programme.[10] The Intercollegiate Surgical Curriculum Programme has set forth a number of workplace-based assessments, including the widely-used procedure-based assessments, which have since been implemented in an effort to improve formative feedback and support learning, as well as to inform summative decisions.

Despite progress in how surgical trainees are assessed, feasibility continues to be of concern, as frequent assessments are time-consuming, costly, and contribute to

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Correspondence: Inquiries to Ranil Sonnadara, Department of Surgery, McMaster University, A. N. Bourns Science Building Room 131, 1280 Main Street West, Hamilton, Ontario L8S 4L8, Canada; fax: (866) 248 - 2972; e-mail: ranil@mcmaster.ca

examiner and examinee burnout.[11,12] Furthermore, the "gold standard" of assessing trainees' technical skills in the operating room (OR) by staff surgeons faces many challenges,[10] including single rater limitations, patient safety concerns, and pressures to increase OR efficiency.[13] There remains a need for more feasible and objective methods of assessing surgical skills.[14,15]

Video-based assessment may allow educators to circumvent some of the barriers of traditional clinical evaluations, as it does not require direct observation of performance.[16,17] This scoping review examines the use of video-based assessment in surgical training, including validity, reliability, benefits, and challenges. We explore its applicability in both low- and high-stakes assessment scenarios, outline its use in clinical and simulated settings, and discuss several novel applications.

METHODS

PRISMA Extension Guidelines for Scoping Reviews[18] and the 2017 Joanna Briggs Institute Methodology for Scoping Reviews[19] were followed. Please refer to Figure 1 for the PRISMA Diagram. The Web of Science database was used to conduct the scoping review using variations on the following topic words: *video-based assessment, video-based feedback,* and *surgical training*. Articles that focused solely on undergraduate medical education, continuing professional development, nonsurgical disciplines, veterinary surgery, or dental surgery were excluded. The search was conducted in June 2018 and revised in March 2019, and included articles published in English during the Web of Science database timespan (1900-2019). In the initial search, 702 nonduplicate articles were identified. Two reviewers (SM and VM) screened the articles and resolved disagreements by consensus. After title screening, 469 articles remained. Articles then underwent abstract and full-text review. Additional publications were identified by title from the reference lists of included articles. The final review included 199 articles. Articles deemed to be of high quality and relevant to the use of video-based assessment in surgical training were included in the manuscript.

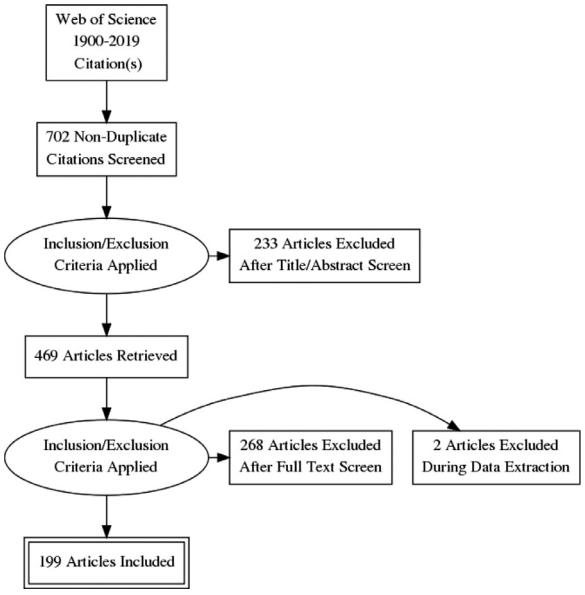

FIGURE 1. PRISMA Diagram summarizing search strategy.

WHY IMPLEMENT VIDEO-BASED ASSESSMENT?

While real-time OR assessments have traditionally been the preferred method for evaluating surgical skills, they are time- and resource-intensive, and require direct faculty observation. Video-based assessment may provide a level of practicality and flexibility that cannot be achieved with traditional evaluations.[20] Many operating suites are now equipped with camera lights and video monitors that can facilitate data collection for video-based assessments, so long as the recordings are of adequate quality, as will be described in this article.[10] Perhaps most notably, the use of video-based techniques in the intraoperative setting may help alleviate both patient safety and OR efficiency concerns introduced by direct assessments of performance, as staff surgeons are able to focus their attention and time on patient care itself, and review trainee performance in detail at a later time.[16] Furthermore, since video-based assessment can be completed remotely when convenient for the examiner, it may help reduce burnout by serving as a more accommodating method.[16] However, to mitigate the risk of burnout, assessors should have protected time for completing evaluations, especially as assessment demands grow with the transition to CBME.[1,2] In addition, the time required to review media may pose a burden to surgical faculty and must be taken into consideration, as described below.

To help mitigate the additional time demands placed on surgical faculty, assessment time may be significantly shortened by fast-forwarding through or editing out nonessential segments of the recordings.[20,21] This is particularly advantageous in the context of CBME where evaluations may only be required for specific components of the procedure. Studies have suggested that editing can reduce assessment time by 50%[20] to 80%[21] without loss of relevant information. However, editing may also introduce bias; an issue which requires further exploration.[22] Who is qualified to carry out video editing remains to be determined. Senior residents may be able to serve as editors and evaluators, which would further

reduce staff physician demand and may help lessen staff burnout;[23] however, there would need to be consideration for the extra burden this would place on residents' time, which arguably might be better spent on patient care, as well as the potential for augmenting resident burnout.

Video-based assessment may provide the advantage of anonymity, allowing for more objective and reliable measures of performance compared with traditional, real-time methods.[23,24] This may be particularly useful for summative evaluations, where objectivity is desired. Of course, the visual and audio information captured in the recordings must be considered if anonymity is desired. Vogt et al. have demonstrated that assessors assign significantly different performance scores depending on whether the videos portray only the trainees' hands, versus trainees from the waist-up.[25] To achieve more objective assessments for summative evaluations, care must be taken to conceal trainees' identities in the recordings. Alternatively, remote evaluators without prior trainee knowledge may be assigned. In these situations, clear procedures must be in place to return scores to trainees in a timely fashion.

Additional benefits of video-based surgical assessment include a potential increase in patient safety and OR efficiency, as staff surgeons are able to maintain their attention and time on patient care and review trainee performance in detail at a later time.[16] Of course, this requires committed educators who are able to dedicate time to review video data and provide feedback in a timely manner after cases. Again, protected time for assessment may be required. Another benefit of video assessment is that evaluating performance via video recordings facilitates assessment by multiple raters, creating a larger pool of evaluations and thus a less biased rating.[26] This allows for reliability to be measured, as multiple assessments of the same performance can be collated and compared. Some studies have demonstrated the efficacy of this approach, with one group demonstrating that average scores from multiple evaluations of edited, deidentified video recordings were able to accurately predict patient outcomes after laparoscopic gastric bypass surgery.[27]

FORMATIVE AND SUMMATIVE ASSESSMENTS

Video-based techniques have been employed to evaluate a wide variety of skills for both formative (typically low-stakes assessment *for* learning) and summative (typically high-stakes assessment *of* learning) purposes.[20,21,26-28] Video-based assessment has been most frequently used to facilitate formative evaluations,[23] which can be a powerful tool for providing feedback to trainees and enhancing the learning process when implemented correctly.[3] Care must be taken to ensure that feedback is delivered to trainees in a timely manner after the skill is performed. To promote specific, personalized, and high-quality feedback, supervisors who know the trainees well should conduct formative assessments. Educators may choose to review the video recordings alongside trainees to provide a coaching-type format of feedback, with the resident providing input as well. Again, this requires dedicated faculty educators, who should be appropriately compensated for time committed to feedback and coaching. Given the need for greater low-stakes assessments under a CBME approach and the potential added convenience of video-based methods, educators may look to increase the use of video-based assessment to help meet increased assessment demands, provide feedback to trainees, and help develop competence.

SIMULATED AND INTRAOPERATIVE SETTINGS

Simulation-based learning and assessment is becoming increasingly popular due to reduced opportunities for teaching in the OR, greater time pressures, and a renewed awareness for patient safety initiatives.[5,7,13] As in the clinical setting, the greater flexibility and anonymity afforded by video-based methods make them an appealing option in simulated settings.[29] Rooney et al.[30] demonstrated that expert coaches could accurately assess performance in the simulation laboratory based upon video recordings and provide constructive feedback, reducing demands on staff surgeons. In some situations, assessment can even be performed by nonexperts; Holst et al. (2015) have shown that the Crowd Sourced Assessment of Technical Skills tool is as effective as expert faculty surgeons in discriminating between varying levels of surgical skill when reviewing video performances of a robotic intracorporeal suturing model.[31] Assessment and feedback delivered by nonexperts can help residents improve their skills through practice in a controlled, risk-free setting, with minimal time and service demands placed on staff surgeons. However, studies comparing nonexpert versus expert feedback and assessment generally involve simple tasks, such as the application of a forearm cast.[32] Further research is needed to assess whether nonexperts can be used as evaluators for more complex procedures.

One of the more common and extensively studied applications of video-based assessment of technical skills in the laboratory uses the Objective Structured Assessment of Technical Skills (OSATS) exam.[21] Video-based

OSATS have been adapted by many surgical specialties and have been shown to be effective for assessing technical skills in the simulated setting.[33] Video-based assessments have been used to evaluate a variety of skills performed on bench models,[33] cadavers,[26] virtual reality simulators,[34] and the Da Vinci Skills robotic simulator.[35] Several simulation-based training courses have started to implement video-based OSATS as well.[36] As resident education continues to rely more heavily on simulation-based training, video-based techniques may similarly find greater applications in facilitating assessment and learning in this setting.

Despite growing opportunities for simulation-based teaching, observation of surgical skills in the intraoperative setting remains the gold standard[10] and may be the best way of capturing what trainees are able to do during clinical practice.[16] Surgeon educators have traditionally valued direct observation of residents' skills to assess autonomy and skill comfort.[37] Observation and feedback in the OR is essential, whether it is provided directly as a skill is performed, at a later time via video recordings, or using the 2 approaches in combination. Video-based assessment has been shown to be effective in determining skill proficiency in the intraoperative environment,[27] and, as mentioned, it may also allow educators to circumvent some of the barriers to assessment in this high-stakes setting as it does not require direct, real-time observation.[13] This lessens patient safety concerns, as supervising surgeons do not have to split their attention between assessing residents' skills and monitoring patient care.[16] Allowing educators to focus entirely on the assessment itself should also lead to higher quality observations and more meaningful feedback for learning.[16] Additionally, OR efficiency is not compromised by a need to survey and document trainees' performance while the procedure is underway. In this way, video-based assessment of intraoperative performance may present a superior alternative to live clinical assessments, as it captures a record of what trainees are able to do during a real intraoperative case, which can later be reviewed by an expert.[23] We, therefore, suggest that educators consider the use of video-based assessment for increasing and enhancing in-training assessments, particularly as the need to document surgical competence becomes increasingly important with the introduction of CBME.

ASSESSMENT TOOLS FOR VIDEO-RECORDED PERFORMANCE

Various tools have been used to rate video-recorded performance. The most common are OSATS-based Global Rating Scales (GRS)[38] and task- and procedure-specific checklists.[20,38-40] These tools have been modified to assess a variety of specific skills, such as the Objective Structured Assessment of Cataract Surgical Skills (OSACSS),[34] the Global Rating Index for Technical Skills (GRITS),[39] condensed versions of the OSATS GRS,[38] the Global Operative Assessment of Laparoscopic Skills (GOALS),[22,41] and the Global Evaluative Assessment of Robotic Skills (GEARS).[42]

New tools are continuously being developed and adapted for formative video-based assessment. It is important for educators to ensure that these tools are able to facilitate meaningful feedback to trainees in order to promote learning and the attainment of competence.[3] To achieve this goal, assessment tools should be procedure-specific, criterion-based, and provide detailed information on performance.[3,43] Educators must strive for evidence-based criterion for tool selection, not only to ensure that performance is accurately assessed, but also to ensure that trainees are provided with the best opportunity to learn surgical skills.

VALIDITY AND RELIABILITY

Ensuring that assessment tools can provide valid and reliable measures of performance are critical when making high-stakes decisions regarding physician competence.[44] Although there are a number of validity frameworks in use with somewhat different approaches, they all view validity as the extent to which the empirical evidence and theoretical rationales support the suitability of inferences reached from test scores.[45] Different forms of validity are sometimes described, including content, criterion, and construct validity, although a unified form of validity is now favored,[46] which may be defined simply as the extent to which an assessment measures what it is purported to measure.[47] Reliability examines how well an assessment is able to repeatedly produce stable and consistent results for equal performances.[47]

As discussed, video recordings may limit rater bias by allowing for anonymity and still achieve a valid measure of performance. However, it has been found that videos only capturing close-up views of the procedure may prevent raters from observing participants' full movements and therefore may actually reduce the validity of the assessment.[25] A balance must be achieved between concealing trainees' identities and capturing enough data for a robust assessment. Alternatively, video recordings may be evaluated by assessors who are unknown to the trainee to limit bias. Other questions must be answered as they may impact the validity of video-based methods. For example, do learners behave differently when they are aware they are being evaluated in-the-moment versus

when they are being video-recorded for later assessment? Are evaluators more critical of video-recorded performance given additional time and resources to focus on the assessment?

There is now a growing body of evidence demonstrating the validity and reliability of video-based assessment tools.[17] In vascular surgery, 2 groups have demonstrated high levels of construct validity for video-based assessment techniques and were able to discriminate between novice and experienced trainees,[16,48] although one of the groups was limited by a small sample size.[16] Beard et al. (2005) assessed several basic surgical skills on simulated models and found good correlations between live and video assessment. Aggarwal et al. (2008) examined 4 types of tools for the video-based assessment of benchmark laparoscopic procedures: an OSATS-based GRS, a condensed version of this scale, a procedure-specific GRS, and a procedure-specific checklist. The OSATS-based GRS and its condensed version demonstrated higher validity and inter-rater reliability than the other 2 assessments. Despite their positive findings, Aggarwal et al. advocate for the use of procedure-specific tools to provide trainees with meaningful feedback and promote learning. Other researchers have developed specific assessment tools for laparoscopic gastric bypass surgery, such as the Bariatric Objective Structured Assessment of Technical Skill (BOSATS) scale, which was found to be both valid and reliable for assessment of intraoperative performances and correlated with the OSATS GRS.[49] Progress has been made with the assessment of nontechnical skills, as well; Spanager et al. (2013) found high inter-rater reliability for the Non-Technical Skills for Surgeons in Denmark (NOTSSdk) tool used in conjunction with video recordings.

It is important to keep in mind that validity and reliability are not properties inherent to the assessment method itself, but are dependent on the context, the individuals being assessed, the assessors, and the tool employed.[50] Although a tool may have been validated for one application of video-based assessment, this does not mean that it will be valid when applied to a new situation. Therefore, if a new tool is to be used or applied to a new context where it will be used to make high-stakes decisions, educators may want to conduct a psychometric analysis of the tool prior to implementation in that setting. This concern is not unique to video-based assessment methods, but to assessment more broadly. Educators must also consider the intended purpose of the tool; if the goal is solely to facilitate formative feedback to trainees and guide the development of competence, demonstrating high levels of accuracy may not be as necessary. Next, we will consider several other factors that may impact the validity and reliability of video-based assessments, including video editing, the quality of the recordings, and rater training.

Video Editing

Shortening recordings is desirable for reducing assessment time. However, this process is often subjective and may constitute a source of bias,[22] which has the potential to reduce validity and reliability.[21,51] Two studies have shown editing to be effective and to maintain the validity of the assessment,[20,27] but there is a general consensus that standards must be developed to delineate optimal editing procedures and reduce editor subjectivity.[21] Standards are also required to guide evaluators in fast-forwarding through videos, depending on the goals of the assessment.[21] Other questions remain: How much time is actually saved when time for editing is considered? Who should be permitted to do the editing, what type of editing is permissible, and how do we ensure proper training of the editors?

Quality of the Recordings

Reliability and validity may be affected by the quality of the recordings, including the frames, angles, and audiovisual information captured.[16] The noisy intraoperative environment, coupled with impedance from face masks, can make capturing audio information difficult. Capturing adequate video coverage for many open abdominal procedures is sometimes a challenge. Video coverage of the operative field may be of lesser concern for laparoscopic surgery, as the video feed is recorded directly off the monitor, capturing the surgeon's exact view.[38] However, a study by Scott et al.[24] did not find strong correlations between global assessment scores of live and video-based assessments when only visual information was recorded from the laparoscope, suggesting that additional audio and visual information from the OR may be necessary in order to obtain accurate assessments. The researchers also noted that video editing employed during the assessment process may have reduced inter-rater reliability, and should be investigated in future studies. Others have shown that the bandwidth can affect reliability, with higher bandwidth (better image quality) improving utility.[41] Technological advances, including wearable devices such as GoPros, may facilitate the collection of audiovisual information and help overcome some of these limitations.[52,53]

Rater Training

As with any type of assessment, there is a recognized need for adequate rater training to ensure the reliability of the assessment.[22] For video-based assessment, different training approaches have been proposed, including mandatory courses on proper tool use, and preparatory

training and practice with the tool in the intraoperative setting.[22] Programs must determine how to optimize faculty training and promote buy-in, especially as the demand for assessment rises and new tools are introduced with the shift to CBME.

LIMITATIONS OF VIDEO-BASED ASSESSMENT

In addition to concerns regarding editing, recording, and rater training, other challenges must be addressed prior to larger-scale or high-stakes implementation of video-based assessment. One issue that is not unique to video-based assessment and remains a challenge for assessment more broadly is the lack of effective assessment strategies for nontechnical skills compared to procedural abilities.[20] Nontechnical skills, such as decision-making, communication, and teamwork, are vital to competent surgical performance yet can be more difficult to evaluate and may also pose an even greater ethical dilemma than recording technical skills alone. Video-based assessment of nontechnical skills may require high-quality audio and visual data capturing a broader intraoperative field. Laparoscopic and arthroscopic video recordings alone may not capture sufficient information about verbal cues, the use of assistants, and other subtle interactions, which are imperative to evaluating nontechnical competencies in the clinical setting.[24] However, as mentioned, technological advances are making the capture and storage of the necessary information more feasible and convenient.

The sensitivity of video-based assessment in distinguishing between trainees with subtle performance differences has been called into question.[54] However, in one study, small differences in practicing bariatric surgeons were differentiated by video-based assessment using a modified OSATS.[27] Nevertheless, until there is sufficient evidence in support of a tool for use in a specific context, video-based assessment may be more appropriate for formative assessments, rather than high-stakes summative evaluations.[54] Additionally, if the goal of assessment is solely to provide feedback and determine when competency has been attained, the need to discriminate between trainees may be reduced, eliminating these concerns.

One must also consider the potential legal implications of using video recordings of procedures for assessment. Some worry that these recordings could open up liability issues should complications arise during a case,[55] although at least one group has reported no instances of medicolegal subpoenas or liability concerns at their institution after 13 years of filming in the emergency room.[55] Obtaining patient consent, ensuring patient confidentiality, and restricting access and storage time of the videos could help reduce legal risks.[55] Similarly, safe data storage and sharing procedures must be put in place, especially when novel technologies or platforms are being employed.

NEW APPLICATIONS

Video-based assessment is being employed for error training, a form of formative assessment in which trainees are video-recorded while practicing and then the recordings are assessed to identify common errors and adverse outcomes.[56] This information can be fed back to trainees to ensure that errors are not repeated in future clinical encounters. Video-based assessment has also found application in the analysis of learning curves[57] and as a way to allow for easier self- and peer-review of performance.[16,58] However, results from studies examining the effectiveness of independent self-review of video recorded performance have been mixed; some have shown that it is effective for enhancing technical surgical skills[59] while others have been unable to show any significant advantages.[60-62] Research from medical education more broadly has shown self-assessment to be inaccurate, with trainees generally over-estimating their abilities.[63] This reinforces the need for expert supervision, assessment, and feedback.[63] The use of video for self-assessment should therefore be cautioned, especially at the early stages of training when learner knowledge is limited. Providing trainees with expert "benchmark" performances to compare with their own performance, or having a supervisor present during the video review process, may be optimal.[64]

Technological advances are improving the efficacy and feasibility of video-based assessment. Novel platforms are being developed to facilitate the distribution and viewing of surgical recordings, which could greatly enhance accessibility and further increase the appeal of video-based methods over real-time assessments.[65] Other programs are being created for viewing the surgeon's hand movements and the operative field simultaneously, using multiple cameras and time-synchronization.[66] This technology can save expert reviewer time and alleviate the concerns regarding insufficient data capture.[66] Although these advances will enhance the utility and feasibility of video-based assessment, novel technologies are often costly. Further cost-benefit analyses are required to determine the added value of these approaches.

CONCLUSIONS

Video-based assessment presents an attractive option for evaluating trainee performance. It may be more objective, convenient, and time-efficient for clinicians than

traditional methods that require direct observation. Video-based assessment is gaining popularity and is being used to evaluate a range of skills in both the OR and simulated settings. There is a growing body of evidence to show that video-based techniques can be valid, reliable, and effective. However, standardized procedures for video recording, editing, and rater training are required. Dedicated surgeon educators who are given protected time and adequate compensation for trainee assessment and feedback are essential. Furthermore, while these techniques may help facilitate formative assessment and feedback, it remains to be determined how to optimally use video-based methods to make summative decisions. Cost-benefit analyses are required to fully understand if the change in outcomes in resident learning is worth the expenses required for these assessments.

Video-based assessment has the potential to become a superior option for evaluating intraoperative performance during surgical training by allowing educators to capture clinical competence while circumventing barriers to assessment in the OR, including concerns for patient safety, intraoperative time-efficiency, and bias from lack of anonymity. The flexibility afforded by video-based assessment could also help ease the burden on supervisors, who are already faced with high rates of burnout. Video-based assessment may be particularly useful in an era of CBME, helping to meet increased assessment demands and provide feedback to trainees to guide the development of competence.

REFERENCES

1. Royal College of Physicians and Surgeons of Canada. Competence by Design: Reshaping Canadian Medical Education. Royal College CBD eBook; 2014. http://www.royalcollege.ca/portal/page/portal/rc/common/documents/educational_initiatives/rc_competency-by-design_ebook_e.pdf. Accessed 1 April 2015.
2. Nasca TJ, Philibert I, Brigham T, Flynn TC. The Next GME accreditation system — rationale and benefits. *N Engl J Med*. 2012;366:1051–1056. https://doi.org/10.1056/NEJMsr1200117.
3. Hattie J, Timperley H. The power of feedback. *Rev Educ Res*. 2007;77:81–112. https://doi.org/10.3102/003465430298487.
4. Future of Medical Education in Canada: Assessment. Royal College White Paper Series. Royal College of Physicians and Surgeons of Canada. http://www.royalcollege.ca/portal/page/portal/rc/common/documents/educational_initiatives/assessment.pdf. Accessed 10 Nov 2013.
5. Sonnadara R, McQueen S, Mironova P, et al. Reflections on current methods for evaluating skills during joint replacement surgery A scoping review. *Bone Joint J*. 2013;95-B:1445–1449. https://doi.org/10.1302/0301-620X.95B11.30732.
6. David MFB. AMEE Medical Education Guide No. 24: Portfolios as a method of student assessment. *Med Teach*. 2001;23:535–551.
7. Ferguson P, Kraemer W, Nousiainen M, et al. Three year experience with an innovative modular competency-based curriculum for orthopaedic training. *Bone Joint J*. 2013;95. e166 1-6.
8. van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34:205–214. https://doi.org/10.3109/0142159X.2012.652239.
9. Pitts D, Rowley DI, Sher JL. Assessment of performance in orthopaedic training. *J Bone Joint Surg Br*. 2005;87-B:1187–1191. https://doi.org/10.1302/0301-620X.87B9.16434.
10. Beard JD. Assessment of surgical skills of trainees in the UK. *Ann R Coll Surg Engl*. 2008;90:282–285. https://doi.org/10.1308/003588408X286017.
11. Wanzel KR, Ward M, Reznick RK. Teaching the surgical craft: from selection to certification. *Curr Probl Surg*. 2002;39:573–659. https://doi.org/10.1067/mog.2002.123481.
12. McQueen S, Petrisor B, Bhandari M, Fahim C, McKinnon V, Sonnadara RR. Examining the barriers to meaningful assessment and feedback in medical training. *Am J Surg*. 2016;211:464–475.
13. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med*. 2006;355:2664–2669.
14. Darzi A, Smith S, Taffinder N. Assessing operative skill. *BMJ*. 1999;318:887–888.
15. Mandel LP, Lentz GM, Goff BA. Teaching and evaluating surgical skills. *Obstet Gynecol*. 2000;95:783–785.
16. Driscoll PJ, Paisley AM, Paterson-Brown S. Video assessment of basic surgical trainees' operative skills. *Am J Surg*. 2008;196:265–272.
17. Scaffidi MA, Grover SC, Carnahan H, et al. A prospective comparison of live and video-based assessments of colonoscopy performance. *Gastroint Endosc*. 2018;87:766–775.

18. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169:467. https://doi.org/10.7326/M18-0850.

19. Peters M, Godfrey C, Khalil H, McInerney P, Parker D, Soares C. Scoping reviews. In: Aromataris E, Munn Z, eds. Joanna Briggs Institute Reviewer's Manual, Adelaide, Australia: Joanna Briggs Inst; 2017.

20. Beard JD, Jolly BC, Newble DI, Thomas WEG, Donnelly J, Southgate LJ. Assessing the technical skills of surgical trainees. *Br J Surg*. 2005;92:778–782. https://doi.org/10.1002/bjs.4951.

21. Dath D, Regehr G, Birch D, et al. Toward reliable operative assessment: the reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surg Endosc*. 2004;18:1800–1804. https://doi.org/10.1007/s00464-003-8157-2.

22. Vassiliou MC, Feldman LS, Fraser SA, et al. Evaluating intraoperative laparoscopic skill: direct observation versus blinded videotaped performances. *Surg Innov*. 2007;14:211–216. https://doi.org/10.1177/1553350607308466.

23. Oestergaard J, Larsen C, Maagaard M, Grantcharov T, Ottesen B, Sorensen J. Can both residents and chief physicians assess surgical skills. *Surg Endosc*. 2012;26:2054–2060. https://doi.org/10.1007/s00464-012-2155-1.

24. Scott DJ, Rege RV, Bergen PC, et al. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. *J Laparoendosc Adv Surg Tech A*. 2000;10:183–190.

25. Vogt VY, Givens VM, Keathley CA, Lipscomb GH, Summitt RL Jr.. Is a resident's score on a videotaped objective structured assessment of technical skills affected by revealing the resident's identity? *Am J Obstet Gynecol*. 2003;189:688–691. https://doi.org/10.1067/S0002-9378(03)00887-1.

26. Laeeq K, Infusino S, Lin SY, et al. Video-based assessment of operative competency in endoscopic sinus surgery. *Am J Rhinol Allergy*. 2010;24:234–237. https://doi.org/10.2500/ajra.2010.24.3434.

27. 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. https://doi.org/10.1056/NEJMsa1300625.

28. Sharma G, Aycart MA, O'Mara L, et al. A cadaveric procedural anatomy simulation course improves video-based assessment of operative performance. *J Surg Res*. 2018;223:64–71.

29. Karam MD, Thomas GW, Koehler DM, et al. Surgical Coaching from Head-Mounted Video in the Training of Fluoroscopically Guided Articular Fracture Surgery. *J Bone Joint Surg Am*. 2015;97:1031–1039.

30. Rooney DM, Hungness ES, DaRosa DA, Pugh CM. Can skills coaches be used to assess resident performance in the skills laboratory? *Surgery*. 2012;151:796–802. https://doi.org/10.1016/j.surg.2012.03.016.

31. Holst D, White L, Brand T, Harper J, Lendvay T. Crowd-sourced assessment of technical skills: an adjunct to urology resident simulation training. *J Endourol*. 2015;29:604–609.

32. Bullock G, Kovacs G, Macdonald K, Story BA. Evaluating procedural skills competence: inter-rater reliability of expert and non-expert observers. *Acad Med*. 1999;74:76–78.

33. Leong JJH, Leff DR, Das A, et al. Validation of orthopaedic bench models for trauma surgery. *J Bone Joint Surg Br*. 2008;90-B:958–965. https://doi.org/10.1302/0301-620X.90B7.20230.

34. Selvander M, Asman P. Ready for OR or not? Human reader supplements Eyesi scoring in cataract surgical skills assessment. *Clin Ophthalmol*. 2013;7:1973–1977. https://doi.org/10.2147/OPTH.S48374.

35. Ramos P, Montez J, Tripp A, Ng CK, Gill IS, Hung AJ. Face, content, construct and concurrent validity of dry lab exercises for robotic training using global assessment tool. *BJU Int*. 2014;113:836–842.

36. Robinson WP, Baril DT, Taha O, et al. Simulation-based training to teach open abdominal aortic aneurysm repair to surgical residents requires dedicated faculty instruction. *J Vasc Surg*. 2013;58:247–253. https://doi.org/10.1016/j.jvs.2013.04.052. .e1-2.

37. Teman NR, Gauger PG, Mullan PB, Tarpley JL, Minter RM. Entrustment of general surgery residents in the operating room: factors contributing to provision of resident autonomy. *J Am Coll Surg*. 2014;219:778–787.

38. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg*. 2008;247:372–379. https://doi.org/10.1097/SLA.0b013e318160b371.

39. Slade Shantz JA, Leiter JR, Collins JB, MacDonald PB. Validation of a global assessment of arthroscopic skills in a cadaveric knee model. *Arthroscopy*. 2013;29:106–112. https://doi.org/10.1016/j.arthro.2012.07.010.

40. Jensen K, Petersen RH, Hansen HJ, Walker W, Pedersen JH, Konge L. A novel assessment tool for evaluating competence in video-assisted thoracoscopic surgery lobectomy. *Surg Endosc*. 2018;32:4173–4182.

41. Choy I, Fecso A, Kwong J, Jackson T, Okrainec A. Remote evaluation of laparoscopic performance using the global operative assessment of laparoscopic skills. *Surg Endosc*. 2013;27:378–383. https://doi.org/10.1007/s00464-012-2456-4.

42. Goh AC, Goldfarb DW, Sander JC, Miles BJ, Dunkin BJ. Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. *J Urol*. 2012;187:247–252. https://doi.org/10.1016/j.juro.2011.09.032.

43. Cantillon P, Sargeant J. Teaching rounds: giving feedback in clinical settings. *BMJ*. 2008;337:1292–1294.

44. Tavares W, Brydges R, Myre P, et al. Applying Kane's validity framework to a simulation based assessment of clinical competence. *Adv Health Sci Educ Theory Pract*. 2018;23:323–338.

45. Messick S. Validity. In: Linn RL, ed. Educational Measurement, Old Tappan, NJ: McMillan; 1989.

46. Kane MT. Current concerns in validity theory. *J Educ Meas*. 2001;38:319–342.

47. Golafshani N. Understanding reliability and validity in qualitative research. *Qual Rep*. 2003;8:597–607.

48. Beard JD. Setting standards for the assessment of operative competence. *Eur J Vasc Endovasc Surg*. 2005;30:215–218.

49. Zevin B, Bonrath EM, Aggarwal R, Dedy NJ, Ahmed N, Grantcharov TP. Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J Am Coll Surg*. 2013;216:955–965. https://doi.org/10.1016/j.jamcollsurg.2013.01.003. .e8.

50. Huang RJ, Limsui D, Triadafilopoulos G. Video-based performance assessment in endoscopy: moving beyond "See One, Do One, Teach One"? *Gastrointest Endosc*. 2018;87:776–777.

51. Datta V, Bann S, Mandalia M, Darzi A. The surgical efficiency score: a feasible, reliable, and valid method of skills assessment. *Am J Surg*. 2006;192:372–378. https://doi.org/10.1016/j.amjsurg.2006.06.001.

52. Paro JAM, Nazareli R, Gurjala A, Berger A, Lee GK. Video-based self-review: comparing google glass and GoPro technologies. *Ann Plast Surg*. 2015;74:S71. https://doi.org/10.1097/SAP.0000000000000423.

53. Kapi E. Surgeon-manipulated live surgery video recording apparatuses: personal experience and review of literature. *Aesthetic Plast Surg*. 2017;41:738–746.

54. van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Stassen LPS, Dankelman J. Objective assessment of technical surgical skills. *Br J Surg*. 2010;97:972–987. https://doi.org/10.1002/bjs.7115.

55. Mackenzie CF, Xiao Y, Hu F-M, Seagull FJ, Fitzgerald M. Video as a tool for improving tracheal intubation tasks for emergency medical and trauma care. *Ann Emerg Med*. 2007;50:436–442. https://doi.org/10.1016/j.annemergmed.2007.06.487.

56. Bonrath EM, Zevin B, Dedy NJ, Grantcharov TP. Error rating tool to identify and analyse technical errors and events in laparoscopic surgery. *Br J Surg*. 2013;100:1080–1088. https://doi.org/10.1002/bjs.9168.

57. van Oldenrijk J, Schafroth MU, Bhandari M, Runne WC, Poolman RW. Time-action analysis (TAA) of the surgical technique implanting the collum femoris preserving (CFP) hip arthroplasty. TAASTIC trial identifying pitfalls during the learning curve of surgeons participating in a subsequent randomized controlled trial (an observational study). *BMC Musculoskelet Disord*. 2008;9:93. https://doi.org/10.1186/1471-2474-9-93.

58. Vaughn CJ, Kim E, O'Sullivan P, et al. Peer video review and feedback improve performance in basic surgical skills. *American J Surg*. 2016;211:355–360.

59. Jamshidi R, LaMasters T, Eisenberg D, Duh QY, Curet M. Video self-assessment augments development of videoscopic suturing skill. *J Am Coll Surg*. 2009;209:622–625.

60. Backstein D, Agnidis Z, Regehr G, Reznick R. The effectiveness of video feedback in the acquisition of orthopedic technical skills. *Am J Surg*. 2004;187:427–432. https://doi.org/10.1016/j.amjsurg.2003.12.011.

61. Backstein D, Agnidis Z, Sadhu R, MacRae H. Effectiveness of repeated video feedback in the acquisition of a surgical technical skill. *Can J Surg*. 2005;48:195–200.

62. Jethwa AR, Perdoni CJ, Kelly EA, Yueh B, Levine SC, Adams ME. Randomized controlled pilot study of video self-assessment for resident mastoidectomy training. *OTO Open*. 2018;2. 2473974X18770417.

63. Hu Y, Tiemann D, Michael Brunt L. Video self-assessment of basic suturing and knot tying skills by novice trainees. *J Surg Educ*. 2013;70:279–283. https://doi.org/10.1016/j.jsurg.2012.10.003.

64. Birnbach DJ, Santos AC, Bourlier RA, et al. The effectiveness of video technology as an adjunct to teach and evaluate epidural anesthesia performance skills. *Anesthesiology*. 2002;96:5–9.

65. Gonzalez R, Martinez JM, Menzo EL, Iglesias AR, Ro CY, Madan AK. Consumer-based technology for distribution of surgical videos for objective evaluation. *Surg Endosc*. 2012;26:2179–2182. https://doi.org/10.1007/s00464-011-2018-1.

66. Jabbour N, Sidman J. Assessing instrument handling and operative consequences simultaneously: a simple method for creating synced multicamera videos for endosurgical or microsurgical skills assessments. *Simul Healthc*. 2011;6:299–303. https://doi.org/10.1097/SIH.0b013e318216fce4.