OPINION

Video review for measuring and improving skill in urological surgery

Zachary J. Prebay, James O. Peabody, David C. Miller and Khurshid R. Ghani

Abstract | Interest is growing within the urological surgery community for objective assessments of technical skill. Surgical video review relies on the use of objective assessment tools to evaluate both global and procedure-specific skill. These evaluations provide structured feedback to surgeons with the aim of improving technique, which has been associated with patient outcomes. Currently, skill assessments can be performed by using expert peer-review, crowdsourcing or computer-based methods. Given the relationship between skill and patient outcomes, surgeons might be required in the future to provide empirical evidence of their technical skill for certification, employment, credentialing and quality improvement. Interventions such as coaching and skills workshops incorporating video review might help surgeons improve their skill, with the ultimate goal of improving patient outcomes.

Global efforts to improve patient care have shifted over the past 5 years towards refining surgical skill instead of the routine monitoring and collection of patient outcomes, given that skill is an important determinant of postoperative outcomes. Many factors, including postoperative care[1] or multidisciplinary teamwork[2], contribute to surgical outcomes, but the influence of a surgeon's technical skill on patient outcomes was largely understudied until it gained interest over the past decade. Traditionally, surgical skill has been inferred by measuring surrogate variables such as surgical case volume. However, these variables might provide an incomplete representation of a surgeon's ability[3]. In 2013, a landmark study of laparoscopic bariatric surgeons empirically demonstrated for the first time what many surgeons have long believed — that a surgeon's superior technical skill, as measured by their peers, was associated with a markedly reduced rate of perioperative complications[4].

The dissemination of minimally invasive surgery, which is easily amenable to video recording, has simplified the process of capturing and assessing technical performance[5]. Video review is a practical method for assessing skill because it enables surgeons to review their performance and reviewers to directly evaluate their peer's performance at a convenient time, eliminating the logistical hurdles associated with in-person evaluations[6]. In contrast to live review, video review enables surgeons to learn from their performance without the additional stress of coordinating the operating room or patient care[6,7]. In addition, reviewers can be blinded to the identity of the surgeon, which is impossible in a live setting and is critical for the unbiased evaluation of skill[8]. Video review might also offer a better assessment of skill than other commonly used methods for evaluating surgical skill, such as robotic simulator performance. Indeed, a 2017 study reported that performance on a da Vinci Skills Simulator did not correlate with the surgical skill of attending surgeons from a variety of disciplines[9]. However, video review must incorporate objective evaluations using validated assessment tools that can discriminate between different levels of skill to ensure its reliability and reproducibility[3,10].

In the future, surgeons could be required to provide empirical evidence of their technical skill for certification, employment, credentialing and quality improvement[11]. However, questions remain regarding the best method for objectively assessing technical skill. Accordingly, postoperative video review might be a solution for directly measuring surgical skill in the modern era. In this Opinion, we explore the role of video review for measuring and assessing the surgical skill of practising urological surgeons and discuss opportunities for quality improvement with the ultimate goal of improving patient care.

Measuring surgical skill

Validated and objective assessment tools are a prerequisite in surgical skill analysis using video review in order to produce reliable assessments. Skill assessment can be conducted by multiple methods (FIG. 1), the most traditional being self-assessment or evaluation by a peer surgeon, although assessment using laypersons recruited from online crowdsourcing has shown promise. To minimize personal bias and human error, one emerging method that completely removes humans from the process is the use of computer-based artificial intelligence to provide objective and timely skill evaluations.

Assessment tools

Validated tools to rate surgical skill were first developed to address the need to assess and standardize the training of surgical residents. Multiple tools have since been developed to evaluate procedure-specific or general ('global') surgical skill, many of which can be used to evaluate skill in either a live or a postoperative setting using video review. One of the first validated assessment tools was the Objective Structured Assessment of Technical Skills (OSATS), a multistation examination for the assessment of technical surgical skills using a global rating scale[12]; although OSATS was designed to assess competency in performing specific surgical tasks, the unit led by Ara Darzi at Imperial College, London, felt that OSATS was best suited for assessing 'simpler tasks' by junior surgeons[13]. With the adoption of minimally invasive surgery, global rating tools such as the Global Operative Assessment of Laparoscopic Skills (GOALS), a five-domain tool to assess general laparoscopic skill using five-point Likert scales, and the Global Evaluative Assessment of Robotic Skills (GEARS), a six-domain tool to assess general

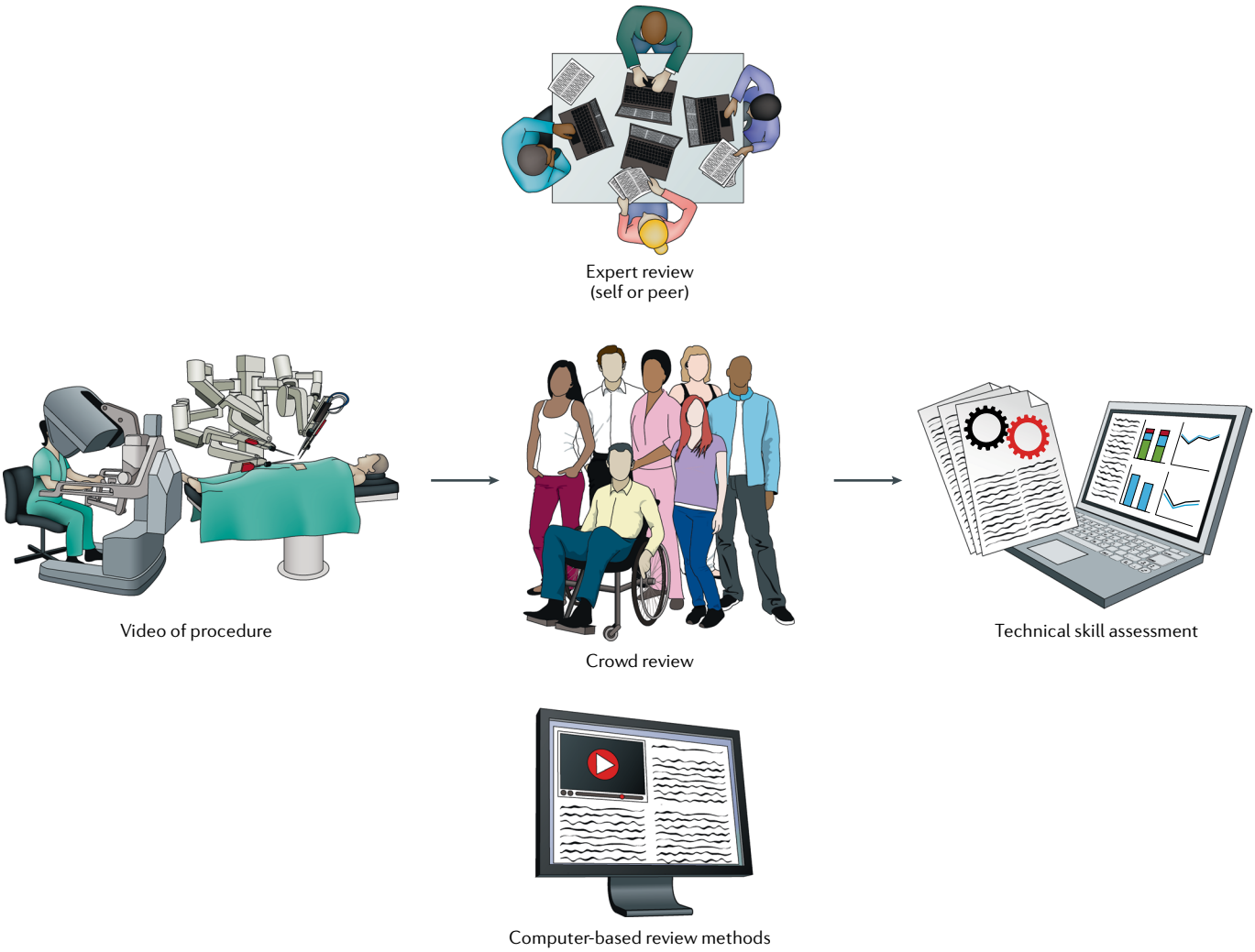

Fig. 1 | **Systematic assessment of surgical skill using video review.** Video review might be a solution for directly assessing surgical skill and can be conducted by multiple methods. By revisiting cases postoperatively with self-assessment, surgeons can appreciate new details of their surgery that they might have otherwise missed in the live operating environment. Peer review is an alternative option for expert review, and surgeons are increasingly accepting the role of peers in their field in evaluating their technical skill. Video review by a peer surgeon can identify areas for improvement that might not be identified during self-assessment. In addition to expert review, crowd review recruits lay individuals through online marketplaces to evaluate surgical videos using globally validated tools. Crowdsourcing is a less costly and faster option than peer review, but the layperson crowd might have a limited ability to evaluate surgical decision-making given their lack of surgical training and knowledge. The use of automated computer-based software to evaluate surgical case videos, which reduces personal bias and human error, is also emerging and could provide a timely and objective method of skill assessment. Further work is needed to develop these systems, and they are a promising application for the future.

robotic skill using five-point Likert scales, were introduced[14,15]. These global assessment tools provide quantitative assessments of a surgeon's overall surgical proficiency in laparoscopic or robotic surgery by evaluating core skills, such as bimanual dexterity, depth perception or efficiency of movement, and can reliably differentiate between the skill levels of practising surgeons and trainees.

By contrast, procedure-specific tools evaluate detailed aspects of a particular operation. During robot-assisted radical prostatectomy (RARP), the urethrovesical anastomosis is one such task that can be evaluated using the Robotic Anastomosis and Competence Evaluation (RACE), a six-domain tool to assess urethrovesical anastomosis skill using a five-point Likert scale[16]. In 2017, urologists from the Michigan Urological Surgery Improvement Collaborative (MUSIC) group, in conjunction with experts in RARP, developed and validated the Prostatectomy Assessment and Competence Evaluation (PACE) tool, a seven-domain tool to assess procedure-specific skill using five-point Likert scales that evaluate the RARP procedure by deconstructing it into seven key parts[17]. Tools such as PACE can provide an additional level of analysis when assessing technical ability and provide structured feedback to surgeons for skill improvement.

Expert review of surgical skill

Evidence shows that having a case video reviewed by another surgeon with content expertise can lead to improved technique and patient outcomes, with self-assessment and assessment by a peer surgeon both showing value.

Self-assessment. Surgeons have long recognized the value of video review for studying variations in technique. Patrick Walsh, a pioneer of open radical prostatectomy (ORP), used self-assessment video review to improve patient outcomes following nerve-sparing ORP[18]. By videotaping ORP in 62 patients and

matching the technique with postoperative erectile function, four operative steps were identified from video review that correlated with outcomes. However, self-assessment of skill has limitations pertaining to bias and accuracy[19], as demonstrated in a video review study of RARP in which intraoperative assessments by the operating surgeon were not reliable predictors of outcomes[20].

Peer assessment. Given that surgeons might have a limited ability to analyse their own performance owing to potential self-bias, surgeons are increasingly accepting the role of external review for skill assessment. Video review has been used as a tool in surgical education for a long time[21], but the use of this process has been recognized as an opportunity for continued skill improvement among practising surgeons, particularly over the past 5 years. In 2005, one of the first studies of video review within a quality assurance programme involving pathologists and surgeons reported that the positive surgical margin rate after laparoscopic radical prostatectomy (LRP) for a single surgeon was reduced from 11% to 4.4% by identifying modifications for neurovascular bundle preservation[22]. In 2015, this concept was demonstrated in a team setting by urologists within a single department, where monthly peer review of RARP videos helped to identify examples of best and suboptimal practice[23].

Crowdsourced review of surgical skill

The use of peer surgeons to evaluate videos can be time consuming and expensive. Owing to these limitations, investigators have explored the role of crowdsourcing to evaluate surgical skill[24,25]. Crowdsourcing refers to the process of obtaining input into a task by using large groups of decentralized, independent individuals to provide aggregated feedback, often online[26]. For assessing surgical videos for skill, crowdsourcing can be conducted through online platforms such as Amazon Mechanical Turk (AMT). The process works by instructing crowdworkers to use validated tools such as GEARS to rate surgical videos, after which they are paid a small sum of money on completion. Most crowdworkers are laypersons with no surgical or medical training. In the first evaluation of this method, ratings by 409 AMT workers for dry-laboratory robotic surgery suture exercises were similar to those obtained by 9 expert robotic surgeons[25]; crowdworkers could evaluate these videos in 5 days, compared with a duration of 24 days for surgeons. Since this initial study, crowdsourcing has been used to evaluate multiple technical scenarios[27,28] and is now available as a commercial platform called Crowd-Sourced Assessment of Technical Skill (C-SATS), a service utilizing expert surgeons or trained reviewers to evaluate surgical videos using validated assessment tools such as GEARS[24].

Urological surgeons in Michigan, working through the MUSIC group, were the first to use crowdsourcing to assess the technical level of surgeons using patient videos[29]. Using de-identified videos of the urethrovesical anastomosis during RARP, they found that peer and crowdsourced ratings of global and procedure-specific skill were strongly correlated. Importantly, both groups agreed on the rank order of the lowest scoring surgeons. Similar findings were also reported by Powers et al.[30] between crowdworkers and expert surgeons when evaluating renal hilar dissection during robot-assisted partial nephrectomy.

Crowdsourcing provides skill ratings at a large scale and in a timely and inexpensive manner, but it is yet to be widely adopted in the surgical community. Limitations to implementation include concerns regarding the use of these skill assessments for decisions on credentialing or re-certification, particularly given that crowdworkers can be laypersons with no surgical training. Furthermore, although a readily available resource, crowdworkers might not be able to evaluate surgical decision-making owing to their lack of surgical training. In a study of ureteroscopy videos, crowdworkers missed obvious ureteral injuries[31]. Despite these limitations, crowdsourcing could potentially serve as a filter for identifying low-scoring surgeons who would then require peer review, summative feedback and interventional coaching strategies to improve performance[29].

Computer analysis of skill

A promising method for skill evaluation is the application of motion analysis and computer vision technology to evaluate the surgeon's hands or instruments in videos and, therefore, determine surgical skill level on the basis of these kinematic assessments. Computer-based methods, if successful, would remove the subjective bias that is inherent in human-based assessments[32]. Motion analysis can be used to track objective metrics such as path length, number of movements and velocity[33]. Previously, these methods relied on the use of onerous devices and sensors[34,35], but now the use of novel software to track regions of interest in videos is feasible[33]. Kinematic data from video-motion analysis can reveal differences in skill level between residents and attending surgeons, such as use of the dominant or nondominant hand[32]. An add-on recording device for the da Vinci robotic system (dvLogger) that can capture video and synchronize it with both kinematic-based and event-based automated performance metrics, such as path length and frequency of third arm swap, was first evaluated by urological surgeons with a wide range of experience[36]. The authors demonstrated both construct and concurrent validity of the kinematic data with concurrent event-based metrics compared with evaluations by expert reviewers using the GEARS tool. A device such as dvLogger could, in the future, function as an operating room 'black box', recording all aspects of the surgery related to skill metrics.

Skill and patient outcomes

Increasing evidence has demonstrated a relationship between video-based peer review of technical performance and patient outcomes (TABLE 1). Surgeons in the Michigan Bariatric Surgery Collaborative (MBSC) were the first to use a systematic approach to study the relationship between skill and surgical outcomes[4]. In this study, 20 bariatric surgeons submitted a single video of a laparoscopic gastric bypass procedure, which was edited to include the most clinically significant and challenging portions. Videos underwent blinded review by at least ten peer surgeons using a modified version of the OSATS and skill ratings were then correlated with adjusted patient outcomes from a clinical registry. The MBSC investigators found that technical skill varied greatly between surgeons and that surgeons in the lowest quartile of skill had substantially increased rates of postoperative complications[4]. However, a subsequent follow-up study by the same group failed to demonstrate a relationship between skill and long-term functional outcomes such as weight loss at 1 year[37]. Additional studies have used video review and determined relationships between surgeons who score highly in techniques such as laparoscopic hernia repair and robot-assisted pancreaticoduodenectomy and reduced rates of hernia recurrence and postoperative pancreatic fistula, respectively[38,39].

In urological surgery, evidence for a relationship between skill and patient outcomes is strongest in the treatment of prostate cancer (TABLE 1). Indeed, the early

Table 1 | **Studies assessing surgical skill and patient outcomes using video-based peer review**

Study	Procedural task	Number of surgeons assessed	Peer reviewers	Tool used	Primary outcomes	Refs
Nonurological surgery						
Arvidsson et al. (2005)	Laparoscopic hernia repair (TAPP and modified Shouldice repair)	25 general surgeons (10 videos of 12 surgeons performing TAPP and 12 in-person evaluations of 13 surgeons performing Shouldice repair)	1 peer surgeon	Unvalidated procedure-specific assessment tool	Significant decrease in hernia recurrence at 5 years (r_s = –0.520; P = 0.019) with surgeons with high performance scores	38
Birkmeyer et al. (2013)	Laparoscopic gastric bypass	20 bariatric surgeons (single video)	33 peer surgeons	OSATS	Compared with surgeons in the top skill quartile, surgeons in the lowest quartile of skill score had significantly increased 30-day complication rates (14.5% versus 5.2%; P < 0.001)	4
Scally et al. (2016)	Laparoscopic gastric bypass	20 bariatric surgeons (single video)	33 peer surgeons	OSATS	No difference between the highest and lowest skill quartiles for the long-term functional outcome of excess body weight loss at 1 year	37
Hogg et al. (2016)	Robotic pancreatico-duodenectomy	133 videos from hepatobiliary surgeons at a single institution	2 hepatobiliary surgeons	Unvalidated procedure-specific assessment tool and modified OSATS	Increased OSATS scores associated with a significant decrease in postoperative pancreatic fistula (P = 0.022)	39
Urological surgery						
Paterson et al. (2016)	Extraperitoneal LRP	1 urological surgeon (200 videos)	2 peer surgeons	Validated procedure-specific assessment tool	Videos of high surgical skill significantly associated with improved urinary continence at 3 months after surgery (r = 0.35, P = 0.013); no correlation between skill and perioperative complications or postoperative erectile function	41
Peabody et al. (2017)	RARP	29 urological surgeons (single video)	56 peer surgeons	GEARS	Surgeons in the top skill quartile had significantly lower rates of excess blood loss (OR = 0.47, P = 0.01) and were associated with fewer events of urethral catheter replacement after its removal (OR = 0.62, P = 0.07) than surgeons in the lowest skill quartile	42
Goldenberg et al. (2017)	RARP	1 urological surgeon (47 videos)	1 peer surgeon	GEARS or GERT	High GEARS score significantly associated with low GERT scores (r = –0.68, P < 0.001); total GEARS score independently predictive of 3-month urinary continence outcomes (OR = 0.55, P = 0.02)	40
Goldenberg et al. (2017)	RARC	3 urological surgeons (9 videos)	5 peer surgeons	5-part questionnaire	No correlation between intraoperative technique and risk of ureteroileal stricture	43

GEARS, Global Evaluative Assessment of Robotic Skills; GERT, Generic Error Rating Tool; LRP, laparoscopic radical prostatectomy; OSATS, Objective Structured Assessment of Technical Skills; RARC, robot-assisted radical cystectomy; RARP, robot-assisted radical prostatectomy; TAPP, transabdominal preperitoneal patch.

return of continence following both RARP[40] and LRP[41] has been correlated with surgical skill as determined by video review. Data presented by the MUSIC group have shown that crowd-based and peer-based evaluations of RARP videos and surgeon skill levels are associated with outcomes such as blood loss, urine leak and re-admission rates[29,42]. However, limitations remain regarding the use of video review. In an evaluation of nine videos of three intraoperative steps of robotic cystectomy performed by three faculty-level surgeons, no association was reported between the perceived risk of ureteroileal stricture (as judged by peer review) and the subsequent development of ureteroileal stricture[43]. Nevertheless, the authors concluded that video review remains

a meaningful pursuit and noted that their study's lack of power and their use of an unvalidated questionnaire might have led to their negative findings.

Applications of video review

Goals of video-based evaluations of surgical skill include quality improvement and hospital credentialing and certification. Quality improvement can occur on multiple levels, such as collaborative video review, individually focused coaching sessions or skills workshops. Video-based evaluation of surgical skill in the credentialing process can assist authorities in assessing surgical competency, particularly for new hires and re-certification.

Collaborative review to improve outcomes

Overall, two studies have demonstrated that collaborative peer review can result in improved patient outcomes[23,44]. Following adoption of an institutional quality assurance programme incorporating peer video review, patient-reported urinary continence at 3 months and potency outcomes at 12 months after RARP were significantly improved[23]. Surgeons used the learning points gained from video review to modify technique, and 3-month patient-reported urinary continence rates improved from 57% to 67% (OR 2.19, $P=0.02$), whereas 12-month erectile function outcomes increased from 21% to 61% potency (HR 3.58, $P=0.04$). Similar strategies were used by Schlomm and colleagues[44] for RARP. Specifically, immediate continence rates after catheter removal were significantly improved by peer review, from 30.9% (no use of pad and no leakage of urine) to 50.1% ($P<0.001$), after one surgeon with superior continence rates was shown to have a different method of apical dissection, which was then shared among the other surgeons[45]. Although this group relied on live intraoperative cross supervision as opposed to video review, they appreciate the role that video review might have in recognizing high performers and sharing techniques among surgeons who are not at the same centre[44].

Surgical coaching

In athletics, elite athletes retain coaches and review videos of their performances, and surgeons could adopt a similar mindset. Teaching from experienced surgeons in conjunction with video review sessions forms the basis of surgical coaching sessions[46], which are currently being used at only a few institutions but have been increasing in popularity with the creation of coaching frameworks over the past 2 years.

Fig. 2 | **Video review for assessing skill and improving technique and outcomes.** The evaluations provided during video review offer surgeons an opportunity to reflect on their ability and identify shortcomings in their current skill level and technique. Surgeons can also use the evaluations to drive personalized peer-to-peer coaching sessions. By working with an experienced surgeon as a coach, surgeons can find additional areas for improvement that would otherwise have been missed with self-review. Additionally, surgeons can benefit from the experience or perspectives of a peer. Ultimately, the goal of skill improvement is to improve patient care. Given that studies have shown that surgical skill correlates with patient outcomes, improving skill will probably improve outcomes, although this notion has yet to be empirically proved.

Various studies have attempted to create coaching frameworks that can serve as a template when institutions adopt coaching programmes[7,47,48]. The Wisconsin Surgical Coaching Framework incorporates three domains of coaching — setting goals, encouraging and motivating, and developing and guiding[7]. Ideal coaches have strong interpersonal skills, are highly experienced and are well respected among their peers[6,7,48].

Coaching for practising surgeons has value both in teaching new skills and in improving current skills. By using postoperative video review, coaching enables a better focus on skill improvement than what is traditionally offered in the operating room alone[46,47]. Coaches can assist surgeons in modifying their technique and help in the adoption of best working practices[49–51]. In a study comparing a peer-coaching programme with conventional training methods, including a recorded web-based instructional video and practice instruments, surgeons who were inexperienced in laparoscopic suturing improved their technical performance, as determined by video review, after receiving individualized coaching and conventional training but not after conventional training alone[52]. Video review and coaching are part of a continuous cycle that includes the assessment of skill and the identification of areas to improve technique, resulting in improved patient outcomes (FIG. 2). The use of surgical coaching could enhance medical education and continued professional development; the current barriers to its widespread adoption are mainly financial, logistical and technical in nature[7].

Skills workshops

An alternative method for promoting surgical skill development is conducting skills workshops. In a study from the early 2000s, urologists practised their laparoscopic skills at one of two ex vivo training laboratories (occurring in 2002 and 2003) followed by expert instruction and videotape review[53]. Following the workshop, surgeons showed a marked improvement in laparoscopic skill metrics. In 2017, surgeons in the MUSIC group demonstrated the feasibility of peer-to-peer video review workshops for RARP, whereby surgeons were paired together on the basis of their skill scores and learning style[54]. In these sessions, surgeons reviewed each other's videos of focused parts of the surgery and provided feedback using a structured format. The workshops helped surgeons to identify potentially beneficial changes in technique that they could implement. Further studies are needed to determine whether workshops or coaching programmes can improve patient outcomes. In addition, whether video review can be used to assess skill improvement over time remains to be studied.

Credentialing and certification

In Japan, a nationwide Endoscopic Surgical Skill Qualification (ESSQ) system has been established using video review to evaluate and certify surgical skill for laparoscopic procedures, including in urological surgery[55,56]. Applicant surgeons submit a representative video that is evaluated by two qualified peer surgeons. Each specialty is graded according to their own assessment guidelines that are based on a scoring system in which surgeons start with a certain number of points based on the procedure being evaluated and lose points for 'dangerous manoeuvres', such as inappropriate dissection of major arteries or inappropriate clip application, with no ability to gain points to offset the loss of points[55]. Long-term follow-up monitoring for certain procedures has shown that certified surgeons have markedly lower complication rates than surgeons who failed to qualify[56]. However, the ESSQ is designed to certify competence in terms of pass or fail and does not use objective tools

to quantify skill[55]. Nevertheless, the ESSQ is the first of its kind and the only large-scale certification system for assessing skill on the basis of peer video review.

Limitations and future directions

A major limitation of peer review is the availability of surgeons to review videos. One solution is to use virtual mentoring through social networking, which was shown to provide timely feedback of robotic surgical procedures[57]. Although crowdsourcing methods enable the quick review of surgical videos, in the future, computer programs that have been validated to assess skill could be an exciting option that is scalable and cost-effective. Computer methods that use convolutional neural networks and machine learning algorithms to differentiate the skill level of surgeons performing robotic surgery are currently under development[58]. In the future, surgeons could conceivably submit a video of a representative or difficult case to a computer program, which could then walk through the case with them and provide automated tutoring. To our knowledge, this application is not currently under development, although proof-of-concept studies have shown the potential of simulation-based intelligent tutoring systems for prostate cryosurgery[59] and robotic surgery education[60]. One potential application of video review that has not yet been studied is its incorporation into a preoperative 'warm-up' programme, similar to routines using ex vivo exercises or robotic simulators that have been previously described[61,62]. Furthermore, if surgical video review becomes a mandatory requirement for credentialing and quality improvement, one unmet challenge that will need addressing is how these videos are stored and accessed. A large-scale system will need to handle the data storage demands in a manner that protects patient confidentiality. Last, key stakeholders will have to determine whether video data are discoverable and which legal safeguards are implemented.

Conclusions

Video review enables surgical skill assessment and its use will probably increase with ongoing refinement by the surgical community. This approach could be of particular interest in urological surgery, which has heavily adopted minimally invasive techniques that enable easy capture of operative video. Video review relies on the use of objective assessment tools to evaluate both global and procedure-specific skill. These evaluations can be used to provide feedback to surgeons with the aim of improving technique and patient outcomes. To date, an accrediting body in Japan is the only such group that uses systematic peer video review to certify surgeons as competent in laparoscopic surgery, including in urological laparoscopy. Peer review remains the standard but is limited by its high cost and the availability of surgeons to provide assessments. One solution is the use of crowdsourcing, which can provide large-scale review of videos and function as a filter to identify surgeons with low skill scores who should undergo a peer-review process. In the future, completely removing humans from the evaluation process through the implementation of machine learning and artificial intelligence to computationally determine skill levels has the potential for a scalable model that might help address the subjective nature of human evaluations. With increasing evidence that technical skill is correlated with patient outcomes, surgical coaching interventions and skills workshops might help surgeons improve their technique and further promote a collaborative learning culture. In urology, the best putative examples of this activity are the use of video review to help surgeons identify changes in technique and skill to improve functional outcomes after robotic prostatectomy. Further work must be performed in refining postoperative video review and surgical coaching frameworks, but the potential of video review to influence and improve patient care makes this a worthwhile endeavour.

*Zachary J. Prebay[1], James O. Peabody[2], David C. Miller[3] and Khurshid R. Ghani[3]**

[1]School of Medicine, Medical College of Wisconsin, Milwaukee, WI, USA.

[2]Center for Outcomes Research, Analytics and Evaluation, Vattikuti Urology Institute, Henry Ford Hospital, Detroit, MI, USA.

[3]Department of Urology, University of Michigan, Ann Arbor, MI, USA.

**e-mail: kghani@med.umich.edu*

https://doi.org/10.1038/s41585-018-0138-2
Published online 8 January 2019

1. Pucher, P. H., Aggarwal, R., Singh, P. & Darzi, A. Enhancing surgical performance outcomes through process-driven care: a systematic review. *World J. Surg.* **38**, 1362–1373 (2014).
2. Epstein, N. E. Multidisciplinary in-hospital teams improve patient outcomes: a review. *Surg. Neurol. Int.* **28**, S295–S303 (2014).
3. Fecso, A. B., Szasz, P., Kerezov, G. & Grantcharov, T. P. The effect of technical performance on patient outcomes in surgery: a systematic review. *Ann. Surg.* **265**, 492–501 (2017).
4. Birkmeyer, J. D. et al. Surgical skill and complication rates after bariatric surgery. *N. Engl. J. Med.* **369**, 1434–1442 (2013).
5. Dimick, J. B. & Varban, O. A. Surgical video analysis: an emerging tool for improving surgeon performance. *BMJ Qual. Saf.* **24**, 490–491 (2015).
6. Hu, Y.-Y. et al. Postgame analysis: using video-based coaching for continuous professional development. *J. Am. Coll. Surg.* **214**, 115–124 (2012).
7. Greenberg, C. C., Ghousseini, H. N., Pavuluri Quamme, S. R., Beasley, H. L. & Wiegmann, D. A. Surgical coaching for individual performance improvement. *Ann. Surg.* **261**, 32–34 (2015).
8. Vassiliou, M. C. et al. Evaluating intraoperative laparoscopic skill: direct observation versus blinded videotaped performances. *Surg. Innov.* **14**, 211–216 (2007).
9. Mills, J. T., Hougen, H. Y., Bitner, D., Krupski, T. L. & Schenkman, N. S. Does robotic surgical simulator performance correlate with surgical skill. *J. Surg. Educ.* **74**, 1052–1056 (2017).
10. Moorthy, K., Munz, Y., Sarker, S. K. & Darzi, A. Objective assessment of technical skills in surgery. *BMJ* **327**, 1032–1037 (2003).
11. Hampton, T. Efforts seek to develop systematic ways to objectively assess surgeons' skills. *JAMA* **313**, 782 (2015).
12. Martin, J. A. et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br. J. Surg.* **84**, 273–278 (1997).
13. Darzi, A., Datta, V. & Mackay, S. The challenge of objective assessment of surgical skill. *Am. J. Surg.* **181**, 484–486 (2001).
14. Vassiliou, M. C. et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am. J. Surg.* **190**, 107–113 (2005).
15. Goh, A. C., Goldfarb, D. W., Sander, J. C., Miles, B. J. & Dunkin, B. J. Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. *J. Urol.* **187**, 247–252 (2012).
16. Raza, S. J. et al. Surgical competency for urethrovesical anastomosis during robot-assisted radical prostatectomy: development and validation of the robotic anastomosis competency evaluation. *Urology* **85**, 27–32 (2015).
17. Hussein, A. A. et al. Development and validation of an objective scoring tool for robot-assisted radical prostatectomy: prostatectomy assessment and competency evaluation. *J. Urol.* **197**, 1237–1244 (2017).
18. Walsh, P. C., Marschke, P., Ricker, D. & Burnett, A. L. Use of intraoperative video documentation to improve sexual function after radical retropubic prostatectomy. *Urology* **55**, 62–67 (2000).
19. Davis, D. A. et al. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* **296**, 1094–1102 (2006).
20. Stern, J. et al. Surgeon perception is not a good predictor of peri-operative outcomes in robot-assisted radical prostatectomy. *J. Robot. Surg.* **5**, 283–288 (2011).
21. Reznick, R. K. Teaching and testing technical skills. *Am. J. Surg.* **165**, 358–361 (1993).
22. Touijer, K. et al. Quality improvement in laparoscopic radical prostatectomy for pT2 prostate cancer: impact of video documentation review on positive surgical margin. *J. Urol.* **173**, 765–768 (2005).
23. Cathcart, P. et al. Achieving quality assurance of prostate cancer surgery during reorganisation of cancer services. *Eur. Urol.* **68**, 22–29 (2015).
24. Lendvay, T. S., White, L. & Kowalewski, T. Crowdsourcing to assess surgical skill. *JAMA Surg.* **150**, 1086–1087 (2015).
25. Chen, C. et al. Crowd-sourced assessment of technical skills: a novel method to evaluate surgical performance. *J. Surg. Res.* **187**, 65–71 (2014).
26. Ranard, B. L. et al. Crowdsourcing—harnessing the masses to advance health and medicine, a systematic review. *J. Gen. Intern. Med.* **29**, 187–203 (2014).
27. Kowalewski, T. M. et al. Crowd-sourced assessment of technical skills for validation of basic laparoscopic urologic skills tasks. *J. Urol.* **195**, 1859–1865 (2016).
28. White, L. W. et al. Crowd-sourced assessment of technical skill: a valid method for discriminating basic robotic surgery skills. *J. Endourol.* **29**, 1295–1301 (2015).
29. Ghani, K. R. et al. Measuring to improve: peer and crowd-sourced assessments of technical skill with robot-assisted radical prostatectomy. *Eur. Urol.* **69**, 547–550 (2016).
30. Powers, M. K. et al. Crowdsourcing assessment of surgeon dissection of renal artery and vein during robotic partial nephrectomy: a novel approach for quantitative assessment of surgical performance. *J. Endourol.* **30**, 447–452 (2016).
31. Kaler, K. S. et al. Crowdsourcing evaluation of ureteroscopic videos using the post-ureteroscopic lesion scale (PULS) to assess ureteral injury. *J. Endourol.* **32**, 275–281 (2017).

32. Frasier, L. L. et al. A marker-less technique for measuring kinematics in the operating room. *Surgery* **160**, 1400–1413 (2016).
33. Glarner, C. E. et al. Quantifying technical skills during open operations using video-based motion analysis. *Surgery* **156**, 729–734 (2014).
34. Datta, V., Chang, A., Mackay, S. & Darzi, A. The relationship between motion analysis and surgical technical assessments. *Am. J. Surg.* **184**, 70–73 (2002).
35. Mason, J. D., Ansell, J., Warren, N. & Torkington, J. Is motion analysis a valid tool for assessing laparoscopic skill? *Surg. Endosc.* **27**, 1468–1477 (2013).
36. Hung, A. et al. Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. *J. Endourol.* **32**, 438–444 (2018).
37. Scally, C. P. et al. Video ratings of surgical skill and late outcomes of bariatric surgery. *JAMA Surg.* **151**, e160428 (2016).
38. Arvidsson, D. et al. Randomized clinical trial comparing 5-year recurrence rate after laparoscopic versus shouldice repair of primary inguinal hernia. *Br. J. Surg.* **92**, 1085–1091 (2005).
39. Hogg, M. E. et al. Grading of surgeon technical performance predicts postoperative pancreatic fistula for pancreaticoduodenectomy independent of patient-related variables. *Ann. Surg.* **264**, 482–491 (2016).
40. Goldenberg, M. G., Goldenberg, L. & Grantcharov, T. P. Surgeon performance predicts early continence after robot-assisted radical prostatectomy. *J. Endourol.* **31**, 858–863 (2017).
41. Paterson, C. et al. Videotaping of surgical procedures and outcomes following extraperitoneal laparoscopic radical prostatectomy for clinically localized prostate cancer. *J. Surg. Oncol.* **114**, 1016–1023 (2016).
42. Peabody, J. O. et al. Surgical skill and patient outcomes after robot-assisted radical prostatectomy. *J. Urol.* **197**, e1129 (2017).
43. Goldenberg, M. G. et al. Feasibility of expert and crowd-sourced review of intraoperative video for quality improvement of intracorporeal urinary diversion during robotic radical cystectomy. *Can. Urol. Assoc. J.* **11**, 331–336 (2017).
44. Schlomm, T., Huland, H. & Graefen, M. Improving outcome of surgical procedures is not possible without adequate quality measurement. *Eur. Urol.* **65**, 1017–1019 (2014).
45. Schlomm, T. et al. Full functional-length urethral sphincter preservation during radical prostatectomy. *Eur. Urol.* **60**, 320–329 (2011).
46. Hu, Y.-Y. et al. Complementing operating room teaching with video-based coaching. *JAMA Surg.* **152**, 318 (2017).
47. Singh, P., Aggarwal, R., Tahir, M., Pucher, P. H. & Darzi, A. A randomized controlled study to evaluate the role of video-based coaching in training laparoscopic skills. *Ann. Surg.* **261**, 862–869 (2015).
48. Stefanidis, D. et al. Developing a coaching mechanism for practicing surgeons. *Surgery* **160**, 536–545 (2016).
49. Ibrahim, A. M., Varban, O. A. & Dimick, J. B. Novel uses of video to accelerate the surgical learning curve. *J. Laparoendosc. Adv. Surg. Tech. A* **26**, 240–242 (2016).
50. Greenberg, C. C., Dombrowski, J. & Dimick, J. B. Video-based surgical coaching: an emerging approach to performance improvement. *JAMA Surg.* **151**, 282–283 (2016).
51. Hogg, M. E. et al. Training in minimally invasive pancreatic resections: a paradigm shift away from "see one, do one, teach one". *HPB* **19**, 234–245 (2017).
52. Palter, V. N., Beyfuss, K. A., Jokhio, A. R., Ryzynski, A. & Ashamalla, S. Peer coaching to teach faculty surgeons an advanced laparoscopic skill: a randomized controlled trial. *Surgery* **160**, 1392–1399 (2016).
53. Nakada, S. Y., Hedican, S. P., Bishoff, J. T., Shichman, S. J. & Wolf, J. S. Expert videotape analysis and critiquing benefit laparoscopic skills training of urologists. *JSLS* **8**, 183–186 (2004).
54. Sarle, R. et al. Surgical skill quality improvement: utilizing a peer video review workshop for surgeons performing robotic prostatectomy. *J. Urol.* **197**, e698 (2017).
55. Matsuda, T. et al. The endoscopic surgical skill qualification system in urological laparoscopy: a novel system in Japan. *J. Urol.* **176**, 2168–2172 (2006).
56. Mori, T., Kimura, T. & Kitajima, M. Skill accreditation system for laparoscopic gastroenterologic surgeons in Japan. *Minim. Invasive Ther. Allied Technol.* **19**, 18–23 (2010).
57. Carter, S. C. et al. Video-based peer feedback through social networking for robotic surgery simulation: a multicenter randomized controlled trial. *Ann. Surg.* **261**, 870–875 (2015).
58. Law, H., Ghani, K. R. & Deng, J. Surgeon technical skill assessment using computer vision based analysis. *Proc. Mach. Learn. Res.* **68**, 88–99 (2017).
59. Sehrawat, A. et al. Simulation-based cryosurgery intelligent tutoring system prototype. *Technol. Cancer Res. Treat.* **15**, 396–407 (2016).
60. Skinner, A. et al. Development and application of a multi-modal task analysis to support intelligent tutoring of complex skills. *Int. J. STEM Educ.* **5**, 14 (2018).
61. Mucksavage, P., Lee, J., Kerbl, D. C., Clayman, R. V. & McDougall, E. M. Preoperative warming up exercises improve laparoscopic operative times in an experienced laparoscopic surgeon. *J. Endourol.* **26**, 765–768 (2012).
62. Moldovano, R., Tărcoveanu, E., Dimofte, G., Lupascu, C. & Bradea, C. Preoperative warm-up using a virtual reality simulator. *JSLS* **15**, 533–538 (2011).

Acknowledgements
The authors thank all the surgeons and support staff from the Michigan Urological Surgery Improvement Collaborative.

Author contributions
K.R.G. and Z.J.P. researched data for the article and wrote the manuscript. D.C.M., K.R.G. and Z.J.P. made substantial contributions to discussion of content. All authors reviewed and/or edited the article before submission.

Competing interests
K.R.G. has received contract support from Blue Cross Blue Shield of Michigan for serving as the Co-Director of the Michigan Urological Surgery Improvement Collaborative. K.R.G. has also received an Intuitive Surgical Inc., Technology Research grant. D.C.M. has received contract support from Blue Cross Blue Shield of Michigan for serving as the Director of the Michigan Urological Surgery Improvement Collaborative. Z.J.P and J.O.P. declare no competing interests.

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

Reviewer information
Nature Reviews Urology thanks M. Goldenberg, R. Satava and P. Pucher for their contribution to the peer review of this work.

RELATED LINKS
AMT: https://www.mturk.com
C-SATS: https://www.csats.com
MUSIC: http://musicurology.com/