



Technical Skills Assessment in Robotic Surgery: A Review of Recent Methods

REVIEW

JACOB L. LAUGHLIN, BS

LIANNE R. JOHNSON, MS

BHARGAV GHANEKAR, MS

MARCIA K. O'MALLEY, PHD

**Author affiliations can be found in the back matter of this article*

HOUSTON
Methodist
DEBAKEY HEART &
VASCULAR CENTER

ABSTRACT

Robot-assisted minimally invasive surgery (RAMIS) offers numerous benefits over traditional open surgery, resulting in greater prevalence of use and range of approved procedures. The proliferation of RAMIS has highlighted a need for effective, robust, and objective methods for assessing robotic surgical skills. Traditionally, assessment has relied on expert observation using structured grading rubrics. Although validated and widely used, this method is also resource intensive and subject to reviewer bias. In response, recent work has explored the potential for more robust assessment methods, including the development of skill-based metrics, crowd-sourced assessment techniques, and automated evaluation systems. This review summarizes recent developments in robotic surgical technical skill assessment, focusing on studies using the da Vinci platform. Assessment methods are grouped into four categories: structured rubrics, skill-based metrics, crowd-sourcing techniques, and automated assessment models. Trends of note include adaptation of established rubrics for specific areas of specialty, the implementation of deep learning models for automated assessment, and a move to integrate crowd-sourcing platforms for efficient and inexpensive evaluation. While traditional grading rubric structures remain the standard, multilevel assessment strategies and objective feedback systems are gaining traction. Future work should seek to integrate task- and movement-based assessment into procedure-level evaluations to create more robust and generalizable models for assessment. These advances show a shift towards data-driven and objective assessment methods, which could improve surgical training and patient outcomes.

CORRESPONDING AUTHOR:

Marcia K. O'Malley, PhD

Department of Mechanical Engineering, Rice University, Houston, Texas, US

omalley@rice.edu

KEYWORDS:

surgical robotics; skill assessment; evaluation methods; machine learning; kinematics

TO CITE THIS ARTICLE:

Laughlin JL, Johnson LR, Ghanekar B, O'Malley MK. Technical Skills Assessment in Robotic Surgery: A Review of Recent Methods. *Methodist DeBaKey Cardiovasc J.* 2025; 21(5):49-59. doi: [10.14797/mdcvj.1667](https://doi.org/10.14797/mdcvj.1667)

INTRODUCTION

Robot-assisted minimally invasive surgery (RAMIS) has become increasingly prevalent for a range of procedures, including proctectomy, cholecystectomy, and hernia repair.¹ The da Vinci system from Intuitive Surgical® is the most common RAMIS platform, installed in hospitals in more than 60 countries and used by over 55,000 trained surgeons.^{2,3} Advantages of RAMIS include 3-dimensional (3D) visualization, improved instrument manipulation, and reduced hospital stays.^{4,5} Despite these advantages, extensive training is needed to gain proficiency on these platforms.⁶ Training curricula typically require surgeons to perform to an acceptable degree, so assessing surgical proficiency is vital when implementing effective robotic surgery training programs. Surgical proficiency is typically characterized by technical skills (eg, instrument handling, collisions, and dexterity), and nontechnical skills (eg, procedural knowledge, stress management, and leadership). Proficiency in both of these categories is important for ensuring desired patient outcomes. Since nontechnical skills are difficult to measure, most assessment methods focus on differentiating based on technical skill scores.⁷

For open, minimally invasive, and robotic procedures, assessment of surgical skill traditionally relies on expert observation of trainee surgical performance. The most widely used assessment methods involve structured grading rubrics—for example, OSATS (Objective Structured Assessment of Technical Skill)—and procedure-specific checklists. These tools aim to provide objective evaluations of surgical ability based on clearly predefined criteria. Such assessments may be performed using physical models such as benchtop simulators, cadaveric specimens, or video recordings of clinical procedures. This diversity in task format allows expert evaluators to assess skills ranging from basic instrument handling to complex full-procedure performance. Despite their widespread use, structured assessments of surgical skill by experts have notable limitations. Primarily, they are time and resource intensive and may suffer from bias or inconsistencies in scoring.⁸

New methods of surgical skill assessment are emerging, particularly for RAMIS, that aim to reduce reliance on the expert evaluators typically required for structured grading assessments. Increasingly, these assessment techniques employ kinematic data from the robotic platform itself or other objective metrics that can be interpreted by nonexpert raters or machine learning models.⁹ Crowd-sourcing approaches have also been explored to lessen evaluator workload, using structured grading rubrics scored by large groups of nonexperts. Furthermore, the integration of robotic kinematic data with advanced computer vision and machine learning techniques is driving the development

of automated surgical skill assessment systems, offering scalable alternatives to expert and crowd-sourced evaluation.¹⁰ To capture the latest developments in RAMIS technical skill assessment that are absent in less recent reviews, we surveyed articles published in the past 10 years, specifically focusing on studies using the da Vinci platform as their primary surgical system. First, we describe our search methodology used to gather the articles included in the survey. Next, we identify and explain the varying approaches to assessment represented in the survey articles. Finally, we discuss the applications and potential gaps in the current assessment literature.

METHODS

SEARCH METHODOLOGY

We used Scopus® to locate full text articles written in English that were published between 2014 and 2024. We used the query (robotic AND surgery) OR (robotic AND assisted AND surgery) AND (technical AND skill) OR (motor AND skill) AND (assessment).

STUDY SELECTION

The initial list of 189 results was screened for relevant articles and duplicates by reviewing the titles and abstracts of each article. The result was a list of 123 unique and potentially relevant articles. We reviewed the reduced list of articles to identify studies that covered topics related to the assessment of surgeons or surgical trainees while using one of the da Vinci robotic surgical platforms developed by Intuitive Surgical. This resulted in a list of 57 articles in the final analysis. Additionally, four papers focusing on automated skill assessment from the references of the curated list were added to create a final list of 61 papers (Figure 1).

RESULTS

To understand the current landscape of assessment of robotic surgical performance, we conducted a review of current literature to identify grading approaches, methodologies, and trends and organized the approaches into categories based on the type of assessment used.

STRUCTURED RUBRICS

Structured rubrics are assessment tools that consist of grading categories with clear criteria for evaluation. The areas assessed typically fall into two categories: technical skills and procedural knowledge. Common examples include Objective Structured Assessment of Technical Skills (OSATS),¹¹ The Global Evaluative Assessment of Robotic

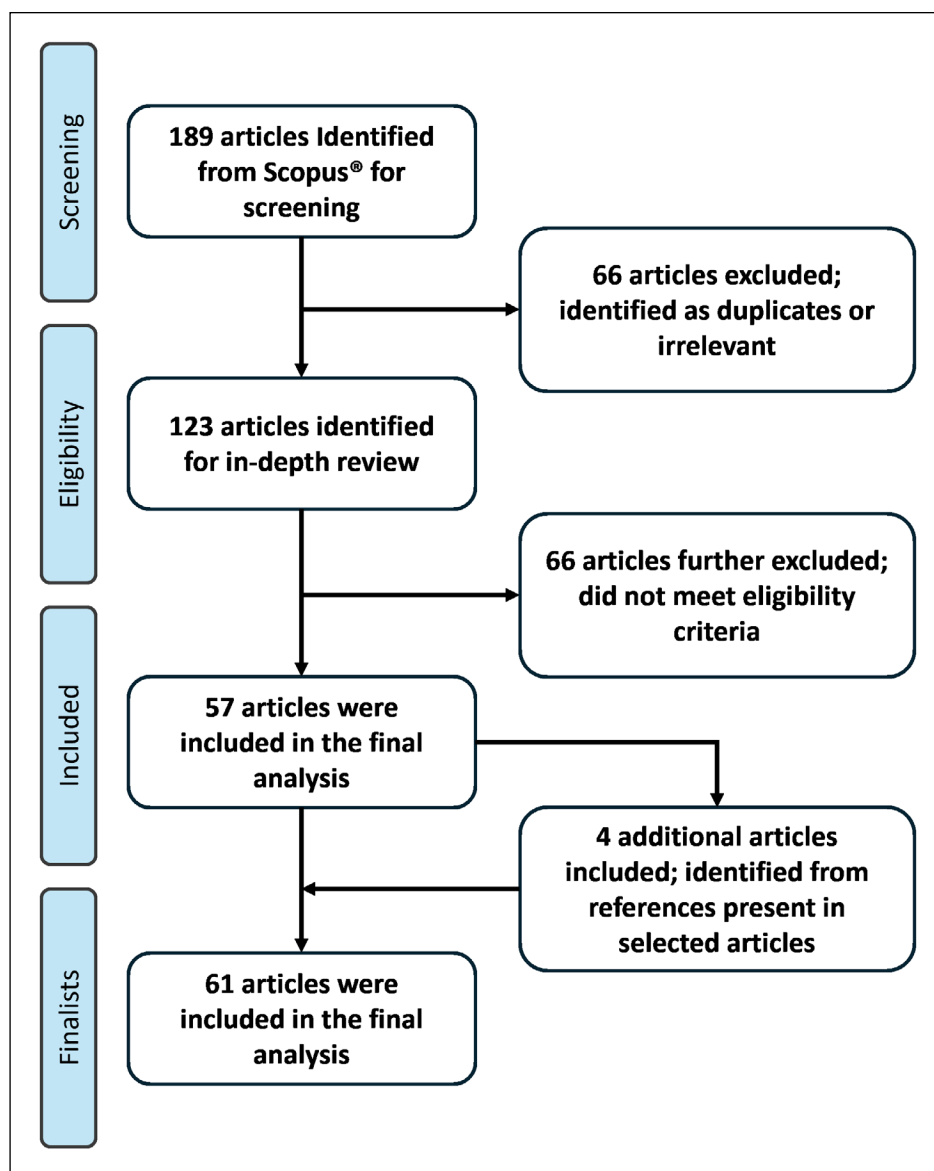


Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram shows selection process for papers included in final analysis.

Skill (GEARS),¹² and the Global Operative Assessment of Laparoscopic Skills (GOALS).¹³ Because these rubrics have been validated to reliably distinguish between different levels of surgical expertise,¹⁴ they are widely regarded as the standard approaches to surgical skill assessment. As a result, structured grading rubrics are commonly used not only for trainee evaluation but also to validate surgical training curricula and emerging assessment methods.

OSATS is a structured rubric built to assess the technical skills of surgical procedures. Initially designed to train technical skills in the laboratory, OSATS has also been used as a performance assessment tool during operations.¹¹ The OSATS rubric consists of seven categories that assess technical skills, such as respect for tissue, time and motion, use of assistants, handling and knowledge of instruments, flow of the operation, and knowledge of the specific

procedure. These categories are assessed on a five-point scale, and the summed value of these individual ratings yields a Global Rating Score (GRS). Due to its established use and extensive validation, OSATS is routinely employed for assessing robotic procedures¹⁵⁻²² or for validating new assessment methods and curricula.²³⁻²⁵

GEARS is a grading rubric used to measure the surgeon's technical performance when performing robotic surgery, evaluating skills in six areas: depth perception, bimanual dexterity, efficiency, force sensitivity, autonomy, and robotic control.¹² Like OSATS, each area is assessed on a five-point scale. GEARS was designed to capture the nuances of operating on a robotic platform, which traditional rubrics like OSATS may overlook. Like OSATS, GEARS is often used to assess surgical procedure performance^{16,19,26-29} or validate other metrics and training curricula.³⁰⁻³⁵

GOALS is a structured grading method that assesses participants on depth perception, bimanual dexterity, efficiency, tissue handling, and autonomy, also based on a five-item global scoring rubric. GOALS has been shown to have high intraclass correlation values for surgeons, showing potential to discriminate between skill levels.¹³

Our survey also identified a number of alternative grading rubrics used to assess robotic surgery, many of which were developed to address specific tasks or limitations of more widely adopted tools such as OSATS and GEARS. These alternative rubrics offer diverse approaches to performance evaluation, often tailored to particular procedures or experimental settings.

A common method among researchers is to adapt standardized rubrics to better fit their specific use case. For example, a modified OSATS (mOSATS) is a rubric adapted from OSATS to include categories for suture handling and quality of the final product and excluding non-applicable categories.³⁶ Other examples, such as robotic OSATS, rate depth perception, tissue handling, dexterity, and efficiency,^{37,38} while anastomosis OSATS, designed for gastrointestinal procedures, evaluates fifteen procedural steps on a three-point scale.³⁹ Another example is the Structured Assessment of Robotic Microsurgical Skills (SARMS), which adapts the Structured Assessment of Microsurgical Skills (SAMS) rubric, with robotic-specific elements like wrist articulation and camera control.⁴⁰ Modified GEARS is another example of a rubric designed to address gaps in the original by adding domains applicable to general robotic skills.^{41,42}

Some rubrics propose new developments instead of simply modifying existing tools. The Assessment of Robotic Console Skills evaluates specific technical skill sets with six evaluation categories such as dexterity, field of view, and force sensitivity, showing reliability in distinguishing across surgeon skill levels.³² Another example is The Objective Clinical Human Reliability Analysis (OCHRA), an error-based tool that attempts to evaluate technical mistakes as the distinguishing factor in skill. Recent adaptations of OCHRA for basic robotic tasks demonstrate strong reliability and correlation with GEARS.³⁵

CROWD-SOURCED ASSESSMENT METHODS

A more recent approach to assessing surgical skills employs crowd-sourcing to generate rubric-based scores. Typically, expert surgeons administer grading rubrics; however, rubric scoring can be a time-intensive and tedious task, making it difficult to recruit experts willing to complete the assessments. Crowd-sourcing approaches solve this issue by having the performance rating done by a large group of laypeople.

Crowd-sourcing of surgical skill assessment is typically facilitated through platforms like Crowd-Sourced

Assessment of Technical Skill (C-SATS), which recruits nonexpert raters, often via services like Amazon Mechanical Turk,⁴³ to score surgical videos using structured grading rubrics such as GEARS in the case of C-SATS. These raters undergo brief training to familiarize them with the rubric criteria before evaluating video clips. Multiple independent ratings are then aggregated using statistical methods to produce reliable scores. This method has been shown to be comparable to expert rater scores when assessing generic technical skills^{37,44} and robotic radical prostatectomy.⁴⁵

SKILL-BASED METRICS

Surgical performance can also be assessed using individual skill-based metrics. These metrics are typically selected because they are inherently valuable in the surgical context (eg, patient outcomes, absence of intraoperative errors) or because they have been separately validated, often by comparison with established grading rubrics.

Completion Time

Time is a prevalent metric for robotic and traditional surgical performance. In the context of surgical assessment, time as a metric can be calculated as the time to complete an entire operation^{33,46} or individual tasks.^{15,20,47-53} Completion time has also been shown to be effective at distinguishing surgical skill level.^{15,16,53} This, paired with its ease of capture, makes completion a desirable and easy metric to include when performing skill assessment.

Kinematic-based Metrics

In robotic surgery, kinematic data refers to motion-related information from surgical instruments and robotic arms, such as position, velocity, acceleration, and joint angles captured from either the surgeon- or patient-side tools. Many robotic platforms are capable of recording this data during procedures; however, access is often limited due to proprietary restrictions.⁵⁴ Despite this, when available, kinematic data allows researchers and system developers to design quantitative metrics for skill assessment and training.

A common approach to assess surgical performance in robotic surgery is to build metrics based on kinematic data collected from the surgical platform. Some of these metrics are based on tool velocity, acceleration, and other higher-order features that aim to capture the surgeon's fine motor control and are used to distinguish technical skills. For example, higher mean tool speeds may indicate either improved efficiency or, conversely, a lack of control in less experienced users.⁵⁵ These metrics can be sensitive to noise and therefore often require filtering or smoothing techniques to ensure robustness.^{51,55}

Another frequently used metric is path length, the total distance traveled by the instrument tips during a

task. In several studies, reductions in path length have been correlated with improved performance on structured rubrics, suggesting construct validity as a measure of surgical expertise.^{47,55,56}

Completion-based Criteria

Researchers have also used task progression to evaluate performance, particularly in motor tasks performed on robotic platforms.^{15,57,58} This method involves breaking down a larger task into smaller sub-steps. Performance scores are typically reported in arbitrary units and are calculated based on how many sub-steps a participant completes. This metric has shown moderate to high correlations to GRS scores.¹⁵

Examples of other criteria include the completeness of tissue closure in suturing, the completeness or quality of tissue resection,⁴⁶ or the requirement of additional resources or steps to complete the task.²¹

Outcome-based Metrics

When looking at outcome-based metrics, surgical skill or expertise is assessed based on the procedure's outcomes. Outcomes include survival rates, intra/postoperative complications (eg, blood loss, tissue injury), and recovery indicators such as length of hospital stay or time to diet resumption.^{21,22,26,58} These outcome measures have been increasingly used to infer technical proficiency across diverse surgical contexts. Additionally, validated assessment tools like OSATS have shown significant associations with some surgical and postoperative complications.²²

Robotic Control

User performance in robotic surgery can be evaluated based on how effectively robotic instruments are manipulated. Key metrics include maintaining instruments within the camera's field of view,³⁴ securely grasping or handling tissue and objects without dropping them,^{23,50} and avoiding collisions with tissue or other instruments.^{34,46} Instrument collisions are commonly measured by frequency, contact duration, or force and vibration levels during interaction.^{16,59}

Studies have shown that experts tend to exhibit lower instrument vibration magnitudes, fewer collisions, and reduced contact forces compared to novices, all of which correlate with higher GRS and GEARS scores.^{16,46}

Precision or Accuracy Errors

Precision or accuracy errors refer to deviations from ideal positions, orientations, or motion paths that characterize expert performance. These may result from poor robotic control or inadequate fine motor skills. Common metrics include needle entry/exit point deviation from predefined targets,⁵⁷ instantaneous or root mean square (RMS) position

and orientation errors,⁶⁰ tool depth error along the viewing axis,⁵⁰ and deviations in grasp position, orientation, or drive path (in or out of plane).⁵⁹ Less common but relevant measures include missed targets, excess tissue piercings, and mechanical or unintentional suture failures.^{21,23,34,59}

Others

Other metrics used to assess performance include the number of gripper activations,⁴⁷ use of additional assist ports,⁶¹ camera use per minute, rate of orientation change, laterality,⁵¹ and other metrics related to knot security—knot tensile strength^{57,58} and the presence of air knots.^{23,34} It is important to note that many metrics—such as camera use per minute, rate of orientation change, laterality, normalized angular displacement—together with mean velocity and spectral arc length were not significantly associated with GEARS.⁵¹

In addition, researchers have used metrics that capture motion economy or efficiency by measuring normalized angular displacement,⁵¹ ribbon area, and console workspace volume.^{47,59}

Other researchers have explored errors more broadly, as any of subjective importance noted by an expert surgeon or as error of omission or commission, with omission-based errors being those behaviors that have occurred partially or not at all and errors of commission being incorrect behaviors that have led to a definable mistake.⁶²

AUTOMATED ASSESSMENT

Analyzing technical skills for robotic surgery in an automated manner can be done by developing machine learning (ML) or deep learning (DL) models that predict scores or skill levels in a supervised manner. Supervised ML/DL models need a training dataset consisting of signals that record relevant robotic tool motion, videos of the tasks being performed, and grading by experts, which serve as the ground truth. Skill assessment could be performed at the task, segment, or gesture level. One early work by Malpani et al. trained support vector machines on certain tool and camera motion feature data present in short segments to create a pairwise preference classifier, obtaining $\geq 85\%$ accuracy when tested on video data from suturing and knot tying tasks.⁴⁷

Most works have developed automated assessment models on the JHU-ISI Gesture and Skill-Assessment Working Set,⁶³ popularly known as the JIGSAWS dataset. JIGSAWS includes synchronized videos and kinematic recordings of dry lab surgical tasks—specifically knot tying, needle passing, and suturing—along with self-reported skill levels and mOSATS scores. Data from the JIGSAWS dataset has been used to train models for predicting skill level and mOSATS scores.^{36,55,64-66} Skill level prediction, typically posed

as a binary (novice/expert) or ternary (novice/intermediate/expert) classification problem, is judged in terms of percentage accuracy (%) in classification. The prediction of mOSATS scores is a regression problem and is typically judged in terms of the Spearman correlation metric (ρ), which is computed between the predicted scores and the expert scores on the test subset.

Prior to the advent of DL methods, research focused on using ML models such as k-nearest neighbors, logistic regression, or support vector machines to predict skill levels or scores^{36,55} and followed a more descriptive analysis approach; for example, they computed features or metrics from kinematic data and trained ML models on the features. Works that subsequently followed typically used raw kinematic tool motion data and trained expressive DL models such as 1-D Convolutional Neural Networks (CNNs)^{54,65} and CNNs with bidirectional long-short term memory (LSTM) models.⁶⁶ These works have demonstrated high precision in predicting skill level.^{36,65} Fawaz et al. achieved 100% accuracy on suturing and needle passing tasks in the JIGSAWS dataset, while Benmansour et al. also showed Spearman correlation values with moderate-to-high correlation ($\rho \geq 0.6$) when comparing predicted scores to OSATS scores.⁶⁶

In addition to supervised approaches, one group was able to use self-supervised learning and uncertainty estimation to adapt a DL model that was originally trained on JIGSAWS data to a novel data set that consisted of videos showing participants performing the ring-transfer task in a virtual reality setting.⁶⁰ The model was able to adapt without the need for manual labeling or task knowledge.

Due to the difficulties present in gathering proprietary kinematic data, some researchers have opted to assess skill through video alone. By using easily accessible video capture in combination with computer vision methods, researchers can estimate kinematic values and perform assessment using the estimates. However, this process can be challenging due to its multidimensional and complex nature. In a recent study by Funke et al.,⁶⁷ 3D CNN layers alongside a Temporal Segment Network were used to classify subject skill level (expert, intermediate, or novice) in the JIGSAWS dataset, achieving $\geq 95\%$ accuracy in all three tasks.⁶⁷ Lee et al. demonstrated that surgical tool tracking from robotic thoracic surgery videos enabled skill level prediction using OSATS and GEARS thresholds with 83% accuracy.⁶⁸ Beyond just performing skill prediction, one group used surgical error annotations on Robot-Assisted Radical Prostatectomy (RARP) operations to train a deep learning model for automated error detection.⁶⁹

Apart from using kinematic motion data or video data, Shafiei et al. asked participants to perform robotic tasks while wearing eye tracking devices and electroencephalogram (EEG) machines.⁷⁰ Using generalized

linear models, the authors demonstrated a link between EEG features, eye gaze features, and experience-related metrics with performance and learning rate evaluation measures, thus highlighting the potential of using auxiliary signals for skill assessment.

While it may be important to automate surgical skill assessment and scoring, it can also be critical to determine whether feedback from such methods can help users effectively learn surgical technical skills. Malpani et al.⁵⁹ conducted a randomized controlled trial and demonstrated that automated real time teaching cues based on grasp orientation during virtual reality-based robotic surgical training are effective. Another recent randomized controlled trial focused on suturing tasks and feedback using artificial intelligence (AI) models.⁷¹ They showed that AI-based automated feedback—provided in the form of video clips demonstrating ideal performance—significantly improved needle handling scores, particularly for under-performers.

DISCUSSION

While a variety of methods exist for assessing surgical training, structured rubrics remain the gold standard. Their widespread adoption is due to strong validation in distinguishing between different levels of surgical expertise as well as their established integration into formal surgical education programs. Rubrics such as OSATS and GEARS are not only used to assess trainee performance but have also become the benchmark for evaluating the effectiveness of alternative assessment techniques and the validity of training curricula. This significant role reinforces their influence in both clinical and research settings, enabling consistent comparisons across studies and institutions.

Most research has assessed surgical skill at a single level, procedure, task, or gesture without integrating these individual assessments into a full performance profile. While some researchers have compared metrics across various levels,⁵⁶ they did not combine them into a full procedure assessment. This highlights an opportunity for future research to develop methods that aggregate lower-level assessments into larger comprehensive evaluations, potentially enabling assessments to be more robust, interpretable, and applicable to actual surgery.

Additionally, we noted that several studies in our survey employed surgical console simulators as platforms for performance assessment. These systems offer practical advantages in that they are low-cost, require no physical setup or expendable resources, and allow for consistent, repeatable task conditions across trials.^{20,24,29,31,49,51,72,73}

However, results are mixed regarding their effectiveness for skill development. Some studies report no significant difference in performance outcomes between simulator-trained and traditionally trained participants⁷³ while other studies suggest that simulator training may result in inferior performance compared to dry lab or cadaver-based training.⁷²

There has also been a strong push towards automation in technical skill assessment, with eight out of thirteen automated skill assessment papers studied in this review published in the last 5 years. Several works have developed models using the JIGSAWS dataset.⁶³ However, the JIGSAWS dataset has several limitations, including a small number of study participants and low-resolution video quality. Thus, there is a need for a large-scale skill assessment dataset that is openly available to the wider research community.

Researchers have also used crowd-sourcing for assessment and validation due to its cost efficiency and quick turnaround times. By implementing platforms such as C-SATS, researchers are able to perform assessments that are shown to correlate with expert raters assessment with lower costs and more quickly than traditional methods. Crowd-sourced grading costs approximately \$16 USD per case compared to \$50 to \$100 USD for expert-based evaluations. This makes for a quick and affordable approach to performing surgical assessment. However, because crowd-sourced raters typically lack specific domain knowledge, much of the assessment is based on visual cues alone. This limits the use of crowd-sourced raters to applications where visual cues are sufficient to determine surgical performance, reducing the overall applicability of this method for medical programs.

CONCLUSION

Robotic surgical skill assessment has evolved significantly in recent years, with structured rubrics such as OSATS and GEARS remaining central due to their validation and widespread adoption. However, growing interest in objective, scalable, and automated methods has driven the development of skill-based metrics, ML models, and vision-based systems. While most studies assess isolated tasks or gestures, there is a clear need for multilevel frameworks that integrate fine-grained and procedural assessments into a unified understanding of surgical competence. As the field moves toward automation and personalized feedback, future efforts should focus on creating large, diverse datasets and cross-platform generalizable models to ensure accurate, efficient, and clinically meaningful evaluation of surgical performance.

KEY POINTS

- Structured rubrics (eg, the Objective Structured Assessment of Technical Skill and the Global Evaluative Assessment of Robotic Skill) remain the most validated and widely used methods but are limited by evaluator time burden and potential bias.
- Crowd-sourced assessment platforms such as Crowd-Sourced Assessment of Technical Skill offer cost-effective and validated alternatives to traditional evaluation methods.
- Automated assessment models using kinematic and video data show high accuracy in predicting skill level of surgeons, particularly with the use of deep learning approaches.
- There is a need for assessment frameworks that integrate more granular assessment into larger performance profiles.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Jacob L. Laughlin, BS  orcid.org/0009-0006-5114-4081

Department of Electrical Engineering, Rice University, Houston, Texas, US

Lianne R. Johnson, MS  orcid.org/0009-0003-4416-0233

Department of Mechanical Engineering, Rice University, Houston, Texas, US

Bhargav Ghanekar, MS  orcid.org/0000-0002-2649-7153

Department of Electrical Engineering, Rice University, Houston, Texas, US

Marcia K. O'Malley, PhD  orcid.org/0000-0002-3563-1051

Department of Mechanical Engineering, Rice University, Houston, Texas, US

REFERENCES

1. **Sheetz KH, Claflin J, Dimick JB.** Trends in the Adoption of Robotic Surgery for Common Surgical Procedures. *JAMA Network Open*. 2020;3(1):e1918911. doi: [10.1001/jamanetworkopen.2019.18911](https://doi.org/10.1001/jamanetworkopen.2019.18911)
2. **Intuitive Surgical [Internet].** Sunnyvale, CA: Intuitive Surgical; c2025. Intuitive reaches 10 million procedures performed using da Vinci surgical systems; 2021 Dec 14 [cited 2025 Jun 27]. Available from: <https://isrg.intuitive.com/news-releases/news-release-details/intuitive-reaches-10-million-procedures-performed-using-da-vinci>

3. **Intuitive Surgical [Internet]**. Sunnyvale, CA: Intuitive Surgical; c2025. Intuitive da Vinci; 2025 [cited 2025 Aug 26]. Available from: <https://www.intuitive.com/en-us/products-and-services/da-vinci>
4. **Corcione F, Esposito C, Cuccurullo D**, et al. Advantages and limits of robot-assisted laparoscopic surgery: preliminary experience. *Surg Endosc*. 2005 Jan;19(1):117-9. doi: [10.1007/s00464-004-9004-9](https://doi.org/10.1007/s00464-004-9004-9)
5. **Han J, Davids J, Ashrafian H, Darzi A, Elson DS, Sodergren M**. A systematic review of robotic surgery: From supervised paradigms to fully autonomous robotic approaches. *Int J Med Robot*. 2022 Apr;18(2):e2358. doi: [10.1002/rcs.2358](https://doi.org/10.1002/rcs.2358)
6. **Barbash GI, Glied SA**. New technology and health care costs — the case of robot-assisted surgery. *New Engl J Med*. 2010 Aug;363(8):701-704. doi: [10.1056/NEJMp1006602](https://doi.org/10.1056/NEJMp1006602)
7. **Birkmeyer JD, Finks JF, O'Reilly A**, et al. Surgical skill and complication rates after bariatric surgery. *New Engl J Med*. 2013 Oct 10;369(15):1434-1442. doi: [10.1056/NEJMs1300625](https://doi.org/10.1056/NEJMs1300625)
8. **Moorthy K, Munz Y, Sarker SK, Darzi A**. Objective assessment of technical skills in surgery. *British Med J*. 2003 Nov 1;327(7422):1032-7. doi: [10.1136/bmj.327.7422.1032](https://doi.org/10.1136/bmj.327.7422.1032)
9. **Levin G, Siedhoff M, Wright KN**, et al. Robotic surgery in obstetrics and gynecology: a bibliometric study. *J Robot Surg*. 2023 Jul 10;17(5):2387-2397. doi: [10.1007/s11701-023-01672-1](https://doi.org/10.1007/s11701-023-01672-1)
10. **Lavanchy JL, Zindel J, Kirtac K**, et al. Automation of surgical skill assessment using a three- stage machine learning algorithm. *Sci Rep*. 2021 Apr 20;11(1):5197. doi: [10.1038/s41598-021-88175-x](https://doi.org/10.1038/s41598-021-88175-x)
11. **Martin JA, Regehr G, Reznick R**, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997 Feb;84(2):273-278. doi: [10.1046/j.1365-2168.1997.02502.x](https://doi.org/10.1046/j.1365-2168.1997.02502.x)
12. **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 Jan;187(1):247-52. doi: [10.1016/j.juro.2011.09.032](https://doi.org/10.1016/j.juro.2011.09.032)
13. **Vassiliou MC, Feldman LS, Andrew CG**, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg*. 2005 Jul;190(1):107-13. doi: [10.1016/j.amjsurg.2005.04.004](https://doi.org/10.1016/j.amjsurg.2005.04.004)
14. **Reiley CE, Lin HC, Yuh DD, Hager GD**. Review of methods for objective surgical skill evaluation. *Surg Endosc*. 2011 Feb;25(2):356-66. doi: [10.1007/s00464-010-1190-z](https://doi.org/10.1007/s00464-010-1190-z)
15. **Tarr M, Rivard C, Petzel A**, et al. Robotic objective structured assessment of technical skills: A randomized multicenter dry laboratory training pilot study. *Female Pelvic Med Reconstr Surg*. 2014 Jul-Aug;20(4):228-36. doi: [10.1097/SPV.000000000000067](https://doi.org/10.1097/SPV.000000000000067)
16. **Gomez E, Aggarwal R, McMahan W, Bark K, Kuchenbecker K**. Objective assessment of robotic surgical skill using instrument contact vibrations. *Surg Endosc*. 2016 Apr;30(4):1419-31. doi: [10.1007/s00464-015-4346-z](https://doi.org/10.1007/s00464-015-4346-z)
17. **Willems J, Shin A, Shin D, Bishop A, Shin A**. A comparison of robotically assisted microsurgery versus manual microsurgery in challenging situations. *Plast Reconstr Surg*. 2016 Apr;137(4):1317-1324. doi: [10.1097/PRS.0000000000002030](https://doi.org/10.1097/PRS.0000000000002030)
18. **Chen Q-Y, Zhong Q, Liu Z-Y**, et al. Surgical outcomes, technical performance, and surgery burden of robotic total gastrectomy for locally advanced gastric cancer a prospective study. *Ann Surg*. 2022 Nov 1;276(5):e434-e443. doi: [10.1097/SLA.0000000000004764](https://doi.org/10.1097/SLA.0000000000004764)
19. **Chen C, Malpani A, Waldram M**, et al. Effect of pre-operative warm-up on trainee intraoperative performance during robot-assisted hysterectomy: a randomized controlled trial. *Int Urogynecol J*. 2023 Nov;34(11):2751-2758. doi: [10.1007/s00192-023-05595-1](https://doi.org/10.1007/s00192-023-05595-1)
20. **Moncayo S, Compagnon R, Caire F**, et al. Transition effects from laparoscopic to robotic surgery skills in small cavities. *J Robot Surg*. 2020 Jun;14(3):525-530. doi: [10.1007/s11701-019-01024-y](https://doi.org/10.1007/s11701-019-01024-y)
21. **Farah E, Abreu A, Rail B**, et al. Adapting to a Robotic Era: The Transferability of Open and Laparoscopic Skills to Robotic Surgery. *J Surg Educ*. 2023 Dec;80(12):1868-1876. doi: [10.1016/j.jsurg.2023.08.014](https://doi.org/10.1016/j.jsurg.2023.08.014)
22. **Liu ZY, Zhong Q, Wang ZB**, et al. Appraisal of surgical outcomes and oncological efficiency of intraoperative adverse events in robotic radical gastrectomy for gastric cancer. *Surg Endosc*. 2024 Apr;38(4):2027-2040. doi: [10.1007/s00464-024-10736-8](https://doi.org/10.1007/s00464-024-10736-8)
23. **Tam V, Zenati M, Novak S**, et al. Robotic Pancreatoduodenectomy Biotissue Curriculum has Validity and Improves Technical Performance for Surgical Oncology Fellows. *J Surg Educ*. 2017 Nov-Dec;74(6):1057-1065. doi: [10.1016/j.jsurg.2017.05.016](https://doi.org/10.1016/j.jsurg.2017.05.016)
24. **Robison W, Patel S, Mehta A**, et al. Can fatigue affect acquisition of new surgical skills? A prospective trial of pre- and post-call general surgery residents using the da Vinci surgical skills simulator. *Surg Endosc*. 2018 Mar;32(3):1389-1396. doi: [10.1007/s00464-017-5820-6](https://doi.org/10.1007/s00464-017-5820-6)
25. **Rail B, Abreu A, Farah E**, et al. Learning Curve of a Robotic Bio-Tissue Intestinal Anastomosis: Implications for Surgical Training Curricula. *J Surg Educ*. 2024 Dec;81(12):103296. doi: [10.1016/j.jsurg.2024.09.015](https://doi.org/10.1016/j.jsurg.2024.09.015)
26. **Shin D, Dalag L, Azhar R**, et al. A novel interface for the telementoring of robotic surgery. *BJU Int*. 2015 Aug;116(2):302-8. doi: [10.1111/bju.12985](https://doi.org/10.1111/bju.12985)
27. **Kelly J, Nash M, Heller N, Lendvay T, Kowalewski T**. Temporal variability of surgical technical skill perception in real robotic surgery. *Int J Comp Assist Radiol Surg*. 2020 Dec;15(12):2101-2107. doi: [10.1007/s11548-020-02253-5](https://doi.org/10.1007/s11548-020-02253-5)

28. **Laca J, Kocielnik R, Nguyen J**, et al. Using Real-time Feedback To Improve Surgical Performance on a Robotic Tissue Dissection Task. *Eur Urol Open Sci*. 2022 Oct 22;46:15-21. doi: [10.1016/j.euros.2022.09.015](https://doi.org/10.1016/j.euros.2022.09.015)
29. **Roberts S, Cen S, Nguyen J**, et al. The Relationship Between Technical Skills, Cognitive Workload, and Errors During Robotic Surgical Exercises. *J Endourol*. 2022 May;36(5):712-720. doi: [10.1089/end.2021.0790](https://doi.org/10.1089/end.2021.0790)
30. **Chowriappa A, Raza S, Fazili A**, et al. Augmented-reality-based skills training for robot-assisted urethrovaginal anastomosis: A multi-institutional randomised controlled trial. *BJU Int*. 2015 Feb;115(2):336-45. doi: [10.1111/bju.12704](https://doi.org/10.1111/bju.12704)
31. **Raison N, Ahmed K, Abe T**, et al. Cognitive training for technical and nontechnical skills in robotic surgery: a randomised controlled trial. *BJU Int*. 2018 Dec;122(6):1075-1081. doi: [10.1111/bju.14376](https://doi.org/10.1111/bju.14376)
32. **Liu M, Purohit S, Mazanetz J, Allen W, Kreaden U, Curet M**. Assessment of Robotic Console Skills (ARCS): construct validity of a novel global rating scale for technical skills in robotically assisted surgery. *Surg Endosc*. 2018 Jan;32(1):526-535. doi: [10.1007/s00464-017-5694-7](https://doi.org/10.1007/s00464-017-5694-7)
33. **Bendre H, Rajender A, Barbosa P, Wason S**. Robotic dismembered pyeloplasty surgical simulation using a 3D-printed silicone-based model: development, face validation and crowdsourced learning outcomes assessment. *J Robot Surg*. 2020 Dec;14(6):897-902. doi: [10.1007/s11701-020-01072-9](https://doi.org/10.1007/s11701-020-01072-9)
34. **Laverty R, Khan M, Patnaik R**, et al. Intentional enterotomies: validation of a novel robotic surgery training exercise. *J Robot Surg*. 2023 Oct;17(5):2109-2115. doi: [10.1007/s11701-023-01625-8](https://doi.org/10.1007/s11701-023-01625-8)
35. **Gorard J, Boal M, Swamynathan V, Ghamrawi W, Francis N**. The application of objective clinical human reliability analysis (OCHRA) in the assessment of basic robotic surgical skills. *Surg Endosc*. 2024 Jan;38(1):116-128. doi: [10.1007/s00464-023-10510-2](https://doi.org/10.1007/s00464-023-10510-2)
36. **Zia A, Essa I**. Automated surgical skill assessment in RAMIS training. *Int J Comput Assist Radiol Surg*. 2018 May;13(5):731-739. doi: [10.1007/s11548-018-1735-5](https://doi.org/10.1007/s11548-018-1735-5)
37. **Newcomb L, Bradley M, Truong T**, et al. Correlation of Virtual Reality Simulation and Dry Lab Robotic Technical Skills. *J Minim Invasive Gynecol*. 2018 May-Jun;25(4):689-696. doi: [10.1016/j.jmig.2017.11.006](https://doi.org/10.1016/j.jmig.2017.11.006)
38. **Nathan A, Patel S, Georgi M**, et al. Virtual classroom proficiency-based progression for robotic surgery training (VROBOT): a randomised, prospective, cross-over, effectiveness study. *J Robot Surg*. 2023 Apr;17(2):629-635. doi: [10.1007/s11701-022-01467-w](https://doi.org/10.1007/s11701-022-01467-w)
39. **Brocke T, Fox C, Clanahan J**, et al. Extrapolative Validity Evidence of the Anastomosis Objective Structured Assessment of Technical Skill (A-OSATS) for Robotic Ileocolic Anastomosis. *J Surg Edu*. 2024 Nov;81(11):1577-1584. doi: [10.1016/j.jsurg.2024.07.021](https://doi.org/10.1016/j.jsurg.2024.07.021)
40. **Selber J, Alrasheed T**. Robotic microsurgical training and evaluation. *Semin Plast Surg*. 2014 Feb;28(1):5-10. doi: [10.1055/s-0034-1368161](https://doi.org/10.1055/s-0034-1368161)
41. **Boal M, Afzal A, Gorard J**, et al. Development and evaluation of a societal core robotic surgery accreditation curriculum for the UK. *J Robot Surg*. 2024 Aug 6;18(1):305. doi: [10.1007/s11701-024-02062-x](https://doi.org/10.1007/s11701-024-02062-x)
42. **Sirajudeen N, Boal M, Anastasiou D**, et al. Deep learning prediction of error and skill in robotic prostatectomy suturing. *Surg Endosc*. 2024 Dec;38(12):7663-7671. doi: [10.1007/s00464-024-11341-5](https://doi.org/10.1007/s00464-024-11341-5)
43. **Amazon Mechanical Turk [Internet]**. Seattle, WA: Amazon Mechanical Turk; c2025. Home page; 2025 [cited 2025 Aug 27]. Available from: <https://www.mturk.com>
44. **White L, Kowalewski T, Dockter R, Comstock B, Hannaford B, Lendvay T**. Crowd-sourced assessment of technical skill: A valid method for discriminating basic robotic surgery skills. *J Endourol*. 2015 Nov;29(11):1295-301. doi: [10.1089/end.2015.0191](https://doi.org/10.1089/end.2015.0191)
45. **Oh P, Chen J, Hatcher D, Djaladat H, Hung A**. Crowdsourced versus expert evaluations of the vesico-urethral anastomosis in the robotic radical prostatectomy: is one superior at discriminating differences in automated performance metrics? *J Robot Surg*. 2018 Dec;12(4):705-711. doi: [10.1007/s11701-018-0814-5](https://doi.org/10.1007/s11701-018-0814-5)
46. **Mendelsohn A, Kim C, Song J**, et al. Transoral Robotic Surgical Proficiency Via Real-Time Tactile Collision Awareness System. *Laryngoscope*. 2020 Dec;130 Suppl 6:S1-S17. doi: [10.1002/lary.29034](https://doi.org/10.1002/lary.29034)
47. **Malpani A, Vedula S, Chen C, Hager G**. A study of crowdsourced segment-level surgical skill assessment using pairwise rankings. *Int J Comput Assist Radiol Surg*. 2015 Sep;10(9):1435-47. doi: [10.1007/s11548-015-1238-6](https://doi.org/10.1007/s11548-015-1238-6)
48. **Shay S, Chrin J, Wang M, Mendelsohn A**. Initial and Long-term Retention of Robotic Technical Skills in an Otolaryngology Residency Program. *Laryngoscope*. 2019 Jun;129(6):1380-1385. doi: [10.1002/lary.27425](https://doi.org/10.1002/lary.27425)
49. **Satava R, Stefanidis D, Levy J**, et al. Proving the effectiveness of the fundamentals of robotic surgery (FRS) skills curriculum: A single-blinded, multispecialty, multi-institutional randomized control trial. *Ann Surg*. 2020 Aug;272(2):384-392. 2020 Aug;272(2):384-392
50. **Mariani A, Pellegrini E, De Momi E**. Skill-Oriented and Performance-Driven Adaptive Curricula for Training in Robot-Assisted Surgery Using Simulators: A Feasibility Study. *IEEE Trans Biomed Eng*. 2021 Feb;68(2):685-694. doi: [10.1109/TBME.2020.3011867](https://doi.org/10.1109/TBME.2020.3011867)
51. **Kelly J, Kowalewski T, Brand T**, et al. Virtual Reality Warm-up Before Robot-assisted Surgery: A Randomized Controlled Trial. *J Surg Res*. 2021 Aug;264:107-116. doi: [10.1016/j.jss.2021.01.037](https://doi.org/10.1016/j.jss.2021.01.037)
52. **Hardon S, Willuth E, Rahimi A**, et al. Crossover-effects in technical skills between laparoscopy and robot-assisted

- surgery. *Surg Endosc*. 2023 Aug;37(8):6015-6024. doi: [10.1007/s00464-023-10045-6](https://doi.org/10.1007/s00464-023-10045-6)
53. **Rahimi A, Hardon S, Willuth E**, et al. Force-based assessment of tissue handling skills in simulation training for robot-assisted surgery. *Surg Endosc*. 2023 Jun;37(6):4414-4420. doi: [10.1007/s00464-023-09905-y](https://doi.org/10.1007/s00464-023-09905-y)
 54. **Cinii Research [Internet]**. Tokyo, Japan: National Institute of Informatics; c2025. DiMaio S, Hasser C. The da Vinci Research Interface; 2025 May 12 [cited 2025 Aug 27]. Available from: <https://cir.nii.ac.jp/crid/1360298341437258240>
 55. **Fard MJ, Ameri S, Darin Ellis R, Chinnam RB, Pandya AK, Klein MD**. Automated robot-assisted surgical skill evaluation: Predictive analytics approach. *Int J Med Robot*. 2018 Feb;14(1). doi: [10.1002/rcs.1850](https://doi.org/10.1002/rcs.1850)
 56. **Vedula S, Malpani A, Ahmidi N, Khudanpur S, Hager G, Chen C**. Task-Level vs. Segment-Level Quantitative Metrics for Surgical Skill Assessment. *J Surg Educ*. 2016 May-Jun;73(3):482-9. doi: [10.1016/j.jsurg.2015.11.009](https://doi.org/10.1016/j.jsurg.2015.11.009)
 57. **Singh H, Modi H, Ranjan S**, et al. Robotic Surgery Improves Technical Performance and Enhances Prefrontal Activation During High Temporal Demand. *Ann Biomed Eng*. 2018 Oct;46(10):1621-1636. doi: [10.1007/s10439-018-2049-z](https://doi.org/10.1007/s10439-018-2049-z)
 58. **Patel R, Suwa Y, Kinross J**, et al. Neuroenhancement of surgeons during robotic suturing. *Surg Endosc*. 2022 Jul;36(7):4803-4814. doi: [10.1007/s00464-021-08823-1](https://doi.org/10.1007/s00464-021-08823-1)
 59. **Malpani A, Vedula S, Lin H, Hager G, Taylor R**. Effect of real-time virtual reality-based teaching cues on learning needle passing for robot-assisted minimally invasive surgery: a randomized controlled trial. *Int J Comput Assist Radiol Surg*. 2020 Jul;15(7):1187-1194. doi: [10.1007/s11548-020-02156-5](https://doi.org/10.1007/s11548-020-02156-5)
 60. **Wang Z, Mariani A, Mencias A, De Momi E, Fey A**. Uncertainty-Aware Self-Supervised Learning for Cross-Domain Technical Skill Assessment in Robot-Assisted Surgery. *IEEE Trans Med Robot Bionics*. 2023 May;5(2):301-11. doi: [10.1109/TMRB.2023.3272008](https://doi.org/10.1109/TMRB.2023.3272008)
 61. **Kheshvadjian M, Nazmifard M, Rawal R**, et al. Implementation and Validation of a High-fidelity Simulation Model for Surgical Resident Training: Management of Acute Intraoperative Hemorrhage during Robot-assisted Surgery. *Urology*. 2024 Sep;191:185-192. doi: [10.1016/j.urology.2024.06.055](https://doi.org/10.1016/j.urology.2024.06.055)
 62. **Wong E, Chu T, Ma R**, et al. Development of a Classification System for Live Surgical Feedback. *JAMA Network Open*. 2023 Jun 1;6(6):e2320702. doi: [10.1001/jamanetworkopen.2023.20702](https://doi.org/10.1001/jamanetworkopen.2023.20702)
 63. **Johns Hopkins University [Internet]**. Baltimore, MD: Johns Hopkins University; c2025. Gao Y, Vedula SS, Reiley CE, et al. Jhu-isi gesture and skill assessment working set (jigsaws): A surgical activity dataset for human motion modeling; 2014 [cited 2025 Aug 27]. Available from: https://cirl.lcsr.jhu.edu/wp-content/uploads/dlm_uploads/2014/09/JIGSAWS.pdf
 64. **Wang Z, Majewicz Fey A**. Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. *Int J Comput Assist Radiol Surg*. 2018 Dec;13(12):1959-1970. doi: [10.1007/s11548-018-1860-1](https://doi.org/10.1007/s11548-018-1860-1)
 65. **Springer Nature [Internet]**. Berlin, Germany: Springer Nature; c2025. Ismail Fawaz H, Forestier G, Weber J, Idoumghar L, Muller P-A. Evaluating Surgical Skills from Kinematic Data Using Convolutional Neural Networks; 2018 Sep 13 [cited 2025 Aug 27]. doi: [10.1007/978-3-030-00937-3_25](https://doi.org/10.1007/978-3-030-00937-3_25)
 66. **Benmansour M, Malti A, Jannin P**. Deep neural network architecture for automated soft surgical skills evaluation using objective structured assessment of technical skills criteria. *Int J Comput Assist Radiol Surg*. 2023 May;18(5):929-937. doi: [10.1007/s11548-022-02827-5](https://doi.org/10.1007/s11548-022-02827-5)
 67. **Funke I, Mees S, Weitz J, Speidel S**. Video-based surgical skill assessment using 3D convolutional neural networks. *Int J Comput Assist Radiol Surg*. 2019 Jul;14(7):1217-1225. doi: [10.1007/s11548-019-01995-1](https://doi.org/10.1007/s11548-019-01995-1)
 68. **Lee D, Yu H, Kwon H, Kong HJ, Lee K, Kim H**. Evaluation of surgical skills during robotic surgery by deep learning-based multiple surgical instrument tracking in training and actual operations. *J Clin Med*. 2020 Jun 23;9(6):1964. doi: [10.3390/jcm9061964](https://doi.org/10.3390/jcm9061964)
 69. **Arxiv [Internet]**. New York, NY: Cornell University; c2025. Psychogyios D, Colleoni E, Van Amsterdam B, et al. SAR-RARP50: Segmentation of surgical instrumentation and action recognition on robot-assisted radical prostatectomy challenge; 2023 Dec 31 [cited 2025 Aug 27]. Available from: <https://arxiv.org/abs/2401.00496>
 70. **Shafiei S, Shadpour S, Intes X, Rahul R, Toussi M, Shafqat A**. Performance and learning rate prediction models development in FLS and RAS surgical tasks using electroencephalogram and eye gaze data and machine learning. *Surg Endosc*. 2023 Nov;37(11):8447-8463. doi: [10.1007/s00464-023-10409-y](https://doi.org/10.1007/s00464-023-10409-y)
 71. **Ma R, Kiyasseh D, Laca J**, et al. Artificial Intelligence-Based Video Feedback to Improve Novice Performance on Robotic Suturing Skills: A Pilot Study. *J Endourol*. 2024 Aug;38(8):884-891. doi: [10.1089/end.2023.0328](https://doi.org/10.1089/end.2023.0328)
 72. **Raison N, Gavazzi A, Abe T, Ahmed K, Dasgupta P**. Virtually Competent: A Comparative Analysis of Virtual Reality and Dry-Lab Robotic Simulation Training. *J Endourol*. 2020 Mar;34(3):379-384. doi: [10.1089/end.2019.0541](https://doi.org/10.1089/end.2019.0541)
 73. **Menso J, Rahimi A, Zwart M**, et al. Robotic hepaticojejunostomy training in novices using robotic simulation and dry-lab suturing (ROSIM): randomized controlled crossover trial. *Surg Endosc*. 2024 Sep;38(9):4906-4915. doi: [10.1007/s00464-024-10914-8](https://doi.org/10.1007/s00464-024-10914-8)

TO CITE THIS ARTICLE:

Laughlin JL, Johnson LR, Ghanekar B, O'Malley MK. Technical Skills Assessment in Robotic Surgery: A Review of Recent Methods. *Methodist DeBakey Cardiovasc J*. 2025;21(5):49-59. doi: [10.14797/mdcvj.1667](https://doi.org/10.14797/mdcvj.1667)

Submitted: 05 July 2025 **Accepted:** 19 August 2025 **Published:** 07 October 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any noncommercial medium, provided the original author and source are credited. See <https://creativecommons.org/licenses/by-nc/4.0/>.

Methodist DeBakey Cardiovascular Journal is a peer-reviewed open access journal published by Houston Methodist DeBakey Heart & Vascular Center.

