Acta Scientific Gastrointestinal Disorders (ISSN: 2582-1091)

Volume 6 Issue 9 September 2023

Research Article

Instruments for the Assessment of Technical and Cognitive Skills in Surgery: A Review Of Scope

Marcus Cezillo*[1] and Renata Bandini Vieira Guglielmi[2]

[1]*Department of Trauma and Acute Care Surgery, Hospital Municipal Antônio Giglio, Brazil*

[2]*Hospital Municipal do Campo Limpo, Brazil*

***Corresponding Author:** Marcus Cezillo, Department of Trauma and Acute Care Surgery, Hospital Municipal Antônio Giglio, Brazil.

DOI: 10.31080/ASGIS.2023.06.0558

Received: June 26, 2023

Published: August 16, 2023

Abstract

The acquisition of competencies in surgery is still the main objective of the professional who seeks independence in a safe and optimized way. However, this task requires the development of assessment tools so that there is quantification and comparison, necessary for the acquisition of competencies. The objective is to review the methods of assessing technical and cognitive skills in general surgery and identify tools and methodologies best applicable in each training context and situation. A scope review was conducted, given the heterogeneous nature of the evidence. A systematic search was performed in the databases MEDLINE, EMBASE, LILACS, Web of Science and Scopus and studies related to the evaluation of technical and cognitive skills in surgery were included after careful selection of the reviewers. With the 21 selected studies it was possible to create a table with information whose observation revealed a great heterogeneity in the evaluation methodology regarding the tool used and comparative method. In addition, most of the time the number of individuals evaluated is low, reflecting the difficulty of conducting this type of clinical research dependent on the commitment of the resident with his teaching and of the preceptor and institution, divided between the care and educational activity organized from the local policy. The limitations that make it difficult to standardize the variables and the comparisons between the different assessment tools are still the subject of study. The review of the current state of the main assessment instruments and their characteristics can provide input for rational and practical use in the face of each reality and need.

Keywords: Instruments; Assessment; Cognitive Skills; Surgery; Scope

Introduction

The training of surgeons committed to the safety and care of patients requires a rigorous evaluation of training, seeking autonomy in the activities involved in the teaching process [1]. This concept, started around 1970, gained strength with the training experiences based on Reliable Professional Activities and with the need for competency-based evaluations [2,3].

In this context there was a change of values, the time spent in a certain activity and the number of executions ceased to be the main parameter of proficiency, giving way to the progressive and gradual demonstration of competences, whose degree of aptitude requires, increasingly, methods of evaluations that are easy to apply in an academic environment with limited human and material resources [4].

Although the competency-based assessment ensures the verification of the behavioral and technical capacity of medical practice, both require distinct and specific evaluation methods [5]. The measurement of behavioral activity by direct observation is

feasible by different preceptors, by the surgeon himself in training and by members of the multidisciplinary team, facilitated by routine coexistence, requiring little of the simulated training. The measurement of technical and cognitive ability requires specific tools used in contexts that are often "not ideal", represented by the patient and their various presentations of diseases. Given this reality, the simulated training and the choice of the ideal assessment tool in each context assume central importance in the consolidation of the teaching program [6].

Without consensus in the literature, different teaching programs use different evaluation methods, such as direct evaluation by specialists (preceptors), using generic evaluation scales or checklist for a given task, computerized analysis through software that quantifies and qualifies a certain task, either in a real or simulated environment [7,8].

Given this context, more important than a standardization of evaluation methods would be the proper application of each method in its necessary context. For this, it is of great importance a compilation on the characterization of each tool, its mode of application and limitations presented.

Methodology

Study protocol

A scope review was carried out due to the scarcity and heterogeneity of studies, providing little evidence in the current literature and generating great difficulty in the search for concrete information that fosters protocols on consistent evaluation methods in each teaching program. Thus, this method of literature review can combine quantitative and qualitative data through the identification of the problem, literature search, evaluation, analysis and presentation of the data found, based on the use of the protocol *Preferred Reporting Items for Systematic Reviews for Scoping Review* (PRISMA-ScR) [9] and reviewed by the team of participating researchers.

Search strategy and source of information

A systematic and comprehensive bibliographic search was performed on July 14, 2022 in the PubMed, Embase, Web of Science, Scopus and LiLACs databases. Through discussion between the team of researchers and an experienced librarian (reviewed by another librarian) the descriptors used in the research were selected (Table 1), based on the *Peer Review of Electronic Search Strategies (PRESS) checklist* [10], seeking to identify potential relevant studies.

In addition, we manually searched the reference lists of relevant articles, and the titles of interest were also included for further evaluation and selection. Searches were also conducted in the "gray literature," including theses, dissertations, and ongoing studies.

Assessment tool	Resident	Surgery	Skill
Assesment	Residency	Surgical	Hability

Tabela 1: Descritores e suas variações usados na pesquisa, composta na search string: [(assesment tool) AND (resident or residency) AND (surgical or surgery) AND (skill OR hability)].

Eligibility criteria

The review included studies published between 2012 and 2022 that presented or discussed methods of evaluation of technical and cognitive skills involving surgeons in training in the experimental models (living tissue or simulator) and also studies that involved evaluation of practical training in surgical procedures related to General Surgery.

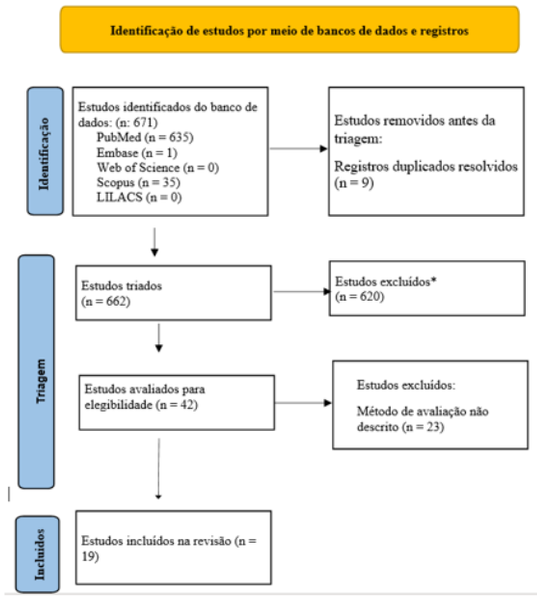
Studies without restriction of language, country of origin or published journal were included and studies that presented training tools without description of the evaluation methodology were not included. Case reports and any study on training in areas other than the surgical area were also excluded.

Selection of studies

All reviewers evaluated the same final number of selected publications, discussed the results and the methodology of screening and data extraction before starting the selection for this review, promoting increased consistency among reviewers. Two independently working reviewers sequentially evaluated the titles, abstracts, and then the full text of all publications identified by our searches for potentially relevant publications. Disagreements between reviewers in the selection of studies and data extraction were resolved by discussion and consensus or further discussion with other reviewers was established, if necessary.

For the selection of the studies, the Rayyan Platform [11] was used, which offers a practical tool and appropriate to the methods and objectives of current research.

Flow chart



*Abstracts excluded because there are no inclusion criteria present (case report, case series, studies without description of the evaluation methodology). Search conducted on 07/14/2022.

Tabulation of data

A table, using the Microsoft Excel application, was developed by both reviewers (Appendix 1) for the purpose of determining and extracting relevant data from the selected studies. The reviewers independently tabulated information such as the author's name, date of publication, tool and form of evaluation employed, and results obtained.

Findings

The studies involving the validation of evaluation tools use the *check-list* of activities and procedures indispensable to performance with autonomy and safety.

This checklist method was applied in person: intraoperatively, preoperatively (breafing) and *postoperatively (*debreafing) of the *selected studies, through the evaluation of the resident's performance in procedures such as cholecystectomy and appendectomy, mainly, in addition to evaluation of cognitive abilities in simulation situations in a face-to-face or non-face-to-face environment.*

Tool name	Environment	Objective of the evaluation	Object of the training	Subject matter of the evaluation
LCAT	Face-to-face and non-face-to-face	Intraoperative	Real case	Technical skill
OSATS	Face-to-face and non-face-to-face	Intraoperative	Real case and simulation	Technical skill
O-SCORE	Face-to-face and non-face-to-face	Intraoperatório	Real case	Technical skill
OMNI	Classroom	Intraoperatório	Real case	Technical skill
CAMEO	Face	Preoperative	Real case	Technical and non-technical skills
OPRS	Face	Intraoperatório	Real case	Technical skills
SIMPL	Face	Intraoperatório	Real case	Technical skills
RRC	Face	Intraoperatório	Real case	Technical skills
NOTSS	Face	Preoperative and intraoperative	Real case	Non-technical skills
GOALS	Face-to-face and non-face-to-face	Intraoperative	Real case and simulation	Technical skills
SPR	Classroom	Preoperative (briefing), intraoperative and postoperative (debriefing)	Real case	Technical and non-technical skills

Table 2: Shows the main characteristics of the selected studies. Real cases can be adapted to simulation scenarios and face-to-face cases can be evaluated, indirectly, through recordings, becoming non-face-to-face, when the evaluation tool allows adaptation.

Discussion

General characteristics of the tools

In the face-to-face modality, there was a report of the use of the *Objective Structured Assessment of Technical Skills* (OSATS), a global scale, widely used that assesses technical competence in open surgeries (intraoperative time). This tool, in addition to being used alone, was also found associated with several other tools such as comparison regarding validity and efficiency. Adapted versions, such as OSAT (Brazilian adaptation of the tool) tested adaptations of the local structure, comparing them with the original tool, resulting in a derivation comparable to the matrix.

Another face-to-face tool reported is the *Visual Analogue Scale* (VAS), adapted to evaluate important moments (steps) of surgery, using a scale ranging from 1 to 7, starting from "very bad" to "excellent" offers a subjective analysis tool, which translates the evaluation of specialists (preceptors) about a given procedure. This tool has often been described as an aid to other instruments, seeking to increase the power of evaluation through the analysis and opinion of experts (preceptors). Another tool for both face-to-face and non-face-to-face use (recorded video analysis) is the Ottawa Surgical Competency Operating Room Evaluation (O-SCORE), a 9-item tool designed to assess technical competence using behavioral anchors. The O-SCORE was able to differentiate the technical level of the surgical resident, with a score equivalent to that of the most validated technical evaluation tool (OSATS) [12,13].

In addition to these checklists, routinely used for open procedures, studies have proposed new specific tools, such as the *Laparoscopic Cholecystectomy-specific Asessment Tool* (LCAT), scoring the execution of decisive steps in a given procedure on a scale of 1 to 4. This new tool, LCAT, [14] was evaluated and validated by comparing the joint evaluation of two other tools: OSATS and VAS. Similar to the LCAT, *The Operative Performance Rating System* (OPRS) assessment instrument [15-17]. allows residents to be classified into 4 to 5 items of specific technical skills and four items that represent general operational skills. The general difficulty of the case and the amount of guidance provided to the resident during the operation is also assessed. These evaluations were validated and widely used to evaluate the intraoperative performance of several different surgical procedures, being quite adaptable.

The "Omni" skills assessment tool suggests better accuracy in gauging technical skill ability as it provides a global assessment across multiple surgical modalities, including open, laparoscopic, robotic, and endoscopic surgery. For this, 5 tasks are evaluated: open intestinal anastomosis, knotting, figure cutting via laparoscopy, handling of the robotic needle and endoscopic bubble burst, in a general score ranging from 0 to 10 for each task [18].

For specific assessment of laparoscopic procedures, the *Global Operative Assessment of Laparoscopic Skills* (GOALS) provides a validated assessment tool. Although it measures domains such as fluidity of movement, it does not define surgical or clinical judgment, for example, autonomy. VASSILOU., *et al.* (2005), considering this, proposed the association of the GOALS with the VAS, allowing a more faithful adaptation of the evaluation to the variation of difficulty of each procedure [19].

In addition to the face-to-face environment, the virtual environment was also used in the technical evaluation of the Surgery Residents. *The Virtual Surgical Patient* (VSP), whose evaluation takes place by application (online format) with instant feedback through the form "*The resident report card*" (RRC) [20], which seeks to simplify the score and language of OSATS through a digital platform with restricted access.

The online (virtual) platform was also the basis of the *System for Improving and Measuring Procedural Learning* (SIMPL) [21], which uses the smartphone as a tool for execution and data storage in an evaluation of a surgical procedure (intraoperative). This form of data acquisition is shown as a positive point for reproduction in a medical residency program, whose medical care merges with the teaching activity, making the evaluation of teaching an activity quite dependent on the preceptor.

There are studies that bring assessment tools not only of technical skills, but also of non-technical skills, such as the validated tool *Clinical Assessment and Management Examination - Outpatient* (CAMEO) [22,23], which evaluates the performance of the resident before the conduct of a clinical case and decision making, having as evaluators the preceptor and the patient, who scores on a scale of 1 to 5 on their perception between "poor" and "excellent". Another tool that also assesses non-technical *skills for surgeons* (NOTSS) [24-26], which observes teamwork and communication skills, skills scored by preceptors at the end of a practical activity. In this

context, the *Surgical Procedure Feedback Rubric* (SPR) [27] offers an evaluation that begins when the patient enters the operating room and finishes when the patient leaves. In addition to technical skills, non-technical skills such as "respect," "knowledge about the patient," and "education" are also "evaluated."

Limitations of the methods

The tools presented in this study may suffer variations in results caused by evaluations performed by different evaluators. However, the greater the number of evaluators of each intern, the lower this bias.

As for the fact that the evaluators already know to which year of medical residency the evaluated person belongs, this may have an influence on the scores. In this case, the non-face-to-face evaluation (recorded procedure) can eliminate this source bias, shielding the evaluators from the identity of the residents. The low adherence on the part of the preceptors can also be a problem regarding the formation of a complete and reliable database. Immediate evaluations, intraoperatively or after (with recordings), are necessary, avoid loss of information and contribute to the *feedback* needed for teaching.

The difference between the previous knowledge of each resident influenced the amount of autonomy practiced by them. Residents with more developed skills had more experience as a surgeon, while residents with less developed skills had more experience as an assistant. For this, tools such as VAS, OSATS, O- SCORE, OPRS consider the level of difficulty of each procedure, seeking to adjust the competence of each resident and their performance to the final score.

Direct observation tools, such as those most commonly used for the evaluation of technical skills, produce a degree of insecurity or differences of opinion among surgeons due to the subjective nature of the evaluation. However, it is generally accepted that experts can identify proficiency when they see it, turning it into a rating scale (also formulated by experts, under the Delphi methodology). Multiple evaluations by multiple observers can minimize this effect.

Due to the validation of the tools, in general, is based on comparisons with other tools already validated and widely used in the literature (OSATS and GOALS), the lack of a group control can make comparison difficult, usually performed between different times in the same individual.

Tools such as GOALS, from the online platform C-SATS, cannot assess the degree of autonomy, due to limitations of analysis of the non-face-to-face format (virtual or recorded video) that do not allow the evaluation of verbal orientation. The evaluation of autonomy is reserved for face-to-face formats only.

There is a study, such as that of HARRIMAN., *et al.* (2009) which recognizes the selection bias caused by the preference of residents in favor of more "malleable" preceptors, despite the mandatory alternation imposed by the stage that by internal agreements can be changed, generating the *"hawk-dove effect*" [28]. Possibly a previous and observed randomization may decrease the chances of negative occurrence of this effect.

In the tools that assess non-technical skills, such as decision making, communication and leadership (such as CAMEO and NOTSS), most of the limitations are due to the difference between the evaluators, prior knowledge and lack of information on how to improve performance in subsequent interactions. This last limitation, related to the *feedback* (*debriefing*), contained in the SPR, for technical evaluations, evaluating only the communication of the non-technical skill part.

Conclusion

Although there is a relatively large number of publications involving training methods in General Surgery, there is a lack of information on consistent evaluation methods. This gap is due, in part, to the difficulty in conducting standardized studies, due to the great variety of characteristics in the heterogeneous groups studied (especially regarding the variable "prior knowledge") and the assessment tools used. This makes it difficult to standardize the variables and to compare the evolution of learning between different moments and between residents.

There is an emerging trend of collective use of these tools, seeking a quick and inexpensive method for assessing technical skills. In this context, the practicality in the collection and storage of data occupies a central role, due to the necessary cooperation of preceptors who divide the time between the care part and the educational part.

Each institution, in view of its reality, can use the tool that best suits the needs and availability of the faculty, associating different methods according to the characteristics of each one, described in this review.

Year	Author	Type of research	Evaluated group	Evaluation method	Objective of the evaluation	Result of the evaluation
2022	Che-vallay., *et al.*	Prospective observational	Residents evaluated live and by recordings of the pro-ceedings	technical evaluation (cholecystectomy check list - LCAT), technical evaluation - OSATS and visual scale for each VAS step)	Assessment tool for laparoscopic cholecystectomy: to assess the inter-ra-ter reliability of this tool.	Live evaluation and recordings have no statistical difference
2020	Campos., *et al.*	Observational prospective	Residents and assis-tants	Residents were filmed and two blinded raters would then evaluate their performance	Validation of a Brazilian Portuguese version of PARTS	OSASTS translated into Portuguese has the same Validity of the original in English
2020	St-Louis	Systematic review + Delphi Analysis	University--affiliated general surgeons, representing a variety of subspecial-ties of Gene-ral Surgery, were invited to answer the ques-tionnaire.	Systematic review + Delphi Analysis	Identificar as habilidade práticas mais importantes (citadas) para deve-lop and validate the OSCAR (objective structured clinical assessment rubric) assessment tool for immediate intra--operative feedback of open technical skills for surgical trainees.	Suturing, tissue and instrument handling, movement economy, instrument knowledge, knot tying, flow, knowledge of procedure, comple-tion time, dissection technique, knowledge of anatomy and sterile technique; 6 of these achieved high or perfect scores and agree-ment after 2 rounds of survey: suturing, sterile technique, knot tying, knowledge of anatomy, knowledge of procedure, and tissue handling.
2019	Harrima., *et al.*	Prospective observational	Urologic trainees - PGY-1 até PGY-5	The resident report card (RRC) is an online, easy-to-use evaluation tool designed to facilitate the creation and distribution of resident technical assessments.	Assesment if a written feedback after each operation can be used to chart surgical progress, can identify under-performing trainees, and will prove bene-ficial for resident learning	The resident report card (RRC) can capture trainee performance over time and against comparator cohorts
2018	Saliken., *et al.*	Prospective observational	Surgical procedures (orthope-dics)	Compare single item evaluation method with 9 item method	Previous work demonstrated that the 9-item O- SCORE can produce valid results; the goal of this study was to de-termine if a single--item performance rating (Is candidate competent to inde-pendently complete procedure: yes or no) completed at a sep- arate viewing would correlate to the O-SCORE, thus increasing feasibility of procedural com-petence assessment.	A single-item performance score correlated highly with the O-S-CORE in an orthopedic setting. A single-item score could be used to supplement a multi-item score with similar Results in orthope-dics.

2018	Steiman., *et al.*	Prospective observational	A total of 10 categorical surgical residents were invited to participate.	Combination of validated assessment tools (multiple choice exam (MCE) questions from the Surgical Council on Resident Education [SCORE], the Clinical Assessment and Management Examination - Outpatient (CAMEO) form, the Virtual Surgical Patient (VSP) website, and a procedure-specific Operative Performance Rating System [OPRS])	Avaliar se 5 Omni tasks, consisting of open bowel anastomosis, knot tying, laparoscopic clover pattern cut, robotic needle drive, and endoscopic bubble pop, were developed by general surgery faculty. Component performance metrics assessed speed, accuracy, and quality, which were scaled into an overall score ranging from 0 to 10 for each task is feasable assesment tool	A single-item performance score correlated highly with the O-SCORE in an orthopedic set- ting. A single-item score could be used to supplement a multi-item score with similar results in orthopedics. There is still benefit in completing multi-item scores such as the O-SCORE evaluations to guide specific areas of improvement and direct feedback.
2018	Cox., *et al.*	Prospective observational	Residents	Apply Omni to medical students and surgical residents	Avaliar se 5 Omni tasks, consisting of open bowel anastomosis, knot tying, laparoscopic clover pattern cut, robotic needle drive, and endoscopic bubble pop, were developed by general surgery faculty. Component performance metrics assessed speed, accuracy, and quality, which were scaled into an overall score ranging from 0 to 10 for each task IS A FEASABLE ASSESMENT TOOL	The Omni holds promise for the evaluation of resident technical skill and early identification of outliers requiring intervention.
2018	Hardon., *et al.*	Prospective observational	Residents	OSATS form and black box with software that evaluates tissue interaction and instrument manipulation	Assess progress before and after training	Questionnaire outcomes indicated that skills and self- confidence improved and that this training should therefore be part of the regular residency training program.

2018	Moore., *et al.*	Prospective observational	(PGY-1) to PGY-5	GS residents postgraduate year 1 (PGY-1) to PGY-5 were anonymously surveyed to evaluate their perceptions of oral and written operative feedback and use of video-based operative resources between 2014 and 2016. Using a 5- point Likert Scale (1 ¼ never, 2 ¼ rarely, 3 ¼ sometimes, 4 ¼ most of the time/frequently, and 5 ¼ always), we evaluated GS residents perceptions of current practices of monthly attending evaluations, intraoperative communicate feedback, and use of video technology. The video was provided to the resident within one week following surgery, thus allowing correlation with the written evaluation. At the conclusion of the video review, the resident completed an assessment of the video- assisted feedback. Again, using the 5-point Likert Scale (1 ¼ not useful, 2 ¼ minimally useful, 3 ¼ moderately useful, 4 ¼ very useful, and 5 ¼ extremely useful), GS residents were surveyed on their perceptions of video-assisted feedback and review. Our surveys were designed ad hoc.	Evaluate perceptions of general surgery residents on video--assisted operative instruction and (2) assess feasibility of providing video review using operative performance rating system to enhance operative feedback during open procedures	The minority of residents (40%) reported that current structured evaluations rarely or never provided meaningful feedback. When feedback was received, 55% residents stated that it was only rarely or sometimes in regard to their operative skills. If feedback on technical skills was received, 46% of residents said that is was never or rarely meaningful or specific; however, 27% believed that feedback would frequently or always result in a change to their operative approach. Receipt of intraoperative or immediate postoperative feedback was reported to occur never or rarely by 32% of GS residents. If it was received, however, 54% of residents stated it was regarding to their technical operative skills. Most of residents (70%) stated they currently watch operations online and most (80%) felt online videos were a helpful educational too lThe video review helped 60% of residents identify unrecognized strengths in the OR, and 80% identify unrecognized weaknesses. The majority of residents (90%) found the video useful for assessing key technical aspects of the case such as dissection of vessels, handling of tissue, and performing the arteriovenous anastomosis. An overwhelming majority (90%) felt that the video review would lead to improved technical skills, wanted to review the video with the attending surgeon for further feedback, and desired expansion of this tool to include additional procedures.

Appendix 1

Year	Author	Type of research	Evaluated group	Evaluation method	Objective of the evaluation	Result of the evaluation
2018	Fryer., *et al*.	Prospective observational	(PG1, n ¼ 6) and categorical (PGY1, n ¼ 5; PGY2, n ¼ 5; PGY3, n ¼ 5; PGY4, n ¼ 5; PGY5, n 1/4 5) general surgery residents	scale of progressive autonomy is a 4-level scale of operative autonomy.7 The first level is "Show and Tell," in which the faculty general surgeon performs critical portions of the operation while explaining the operation and thought process to the surgical resident, who is essentially observing and assisting, but may participate on noncritical portions. The second level is "Active Help," in which the resident performs some portions of the operation, with active guidance from the faculty, who may take over as needed for difficult portions of the operation. The third level is "Passive Help," in which the resident is performing most critical portions of the operation with the faculty assisting but providing guidance only when requested or required for patient safety. The final level is "Supervision Only," in which the resident is the primary surgeon who safely performs most critical portions of the procedure with help from junior resident or OR staff as first assistant. The faculty silently supervises, but does not assist surgically nor provide significant guidance unless requested by the resident or required for patient safety. Ao final residentes respondem questionario de satisfação.	To determine whether the use of PASS (martphone-based system scale for assessing the perform of residents) would not negatively affect resident or faculty satisfaction in the operating room (OR) nor increase mean OR times for cases performed together by residents and faculty	Our data suggest that PASS does no increase mean OR times for the most commonly performed procedures. Resident OR satisfaction did not significantly change during PASS implementation, whereas some changes in faculty satisfaction were noted suggesting that PASS implementation may have had some negative effect with them. Although the effect on faculty satisfaction clearly requires further investigation, our findings support that use of an autonomy-based OR performance assessment system such as PASS does not appear to have a major negative influence on OR times nor OR satisfaction

2016	MacEwan., *et al*.	Prospective observational	Residents and assistants	written questionnaire + Simulated situation in recorded orthopedics assessed using OSATS + O- SCORE	validate O-SCORE results	O-SCORE results differentiated the level of the assessed, agreement between the evaluators, equivalence to the OSATS result. Accuracy and reproducible results
2016	Islam., *et al*.	Prospective observational	Interns, residents and assistants	Video-based assessment instrument for minimally invasive surgery - web based tool - system capable of giving real-time feedback and scoring at the end. Able to evaluate the evolution of surgeons by uploading training videos.	Create a virtual instrument that can give feedback in real time and monitor the development and acquisition of technical skills	The virtual instrument proved to be valid with better evolution in the evaluated who received the feedback in real time and evaluation at the end of the procedure
2016	Bohnen., *et al*.	Prospective observational	General surgery Residents	Provides a 3-question performance assessment for both trainees (self-assessment) and attendings (trainee assessment) following any procedure that they perform together. Attendings additionally can provide specific dictated feedback with every assessment. Data created via SIMPL are available in real time to residents	smartphone-based tool, SIMPL (System for Improving and Measuring Procedural Learning), to make real-time intraoperative performance assessment feasible for every case in which surgical trainees participate, and hypothesized that SIMPL could be feasibly integrated into surgical training programs.	SIMPL can be feasibly integrated into surgical training programs to enhance the frequency and timeliness of intraoperative performance assessment
2016	Shaharan., *et al*.	Prospective observational	General surgery residents at 1st month and at 5th month	First-year surgical trainees were recruited in their first month of the training program. The subjects performed hand knot tying on a bench model. The skill was assessed at baseline in the first month of training and at 5 months. The assessment tools were the Patriot electromagnetic tracking system and Objective Structured Assessment of Technical Skills (OSATS). The trainees' scores were compared to the proficiency score. The data were analyzed using paired t-test and Pearson correlation analysis.	Establish the predictive and concurrent validity of thr "Patriot" as an assessment tool for knot trying and determine the skill retention in first-year surgical trainees after 5 months of training.	The time taken to complete the task and the path length (PL) were significantly shorter at 5 months. In all, 50% of trainees reached the proficiency PL at baseline and at Month 5. Among them, 3 trainees improved their PL to reach proficiency and the other 3 trainees failed to reach proficiency.

2016	Day., *et al*.	Prospective observational	RESIDENTS AND FELLOWS	Recorded evaluation of suture in open surgery - simulation (termino--terminal vascular), 2 evaluators	To evaluate the feasibility, reliability and validity of a new tool for open surgical skills assessment, the 8-minute suture test (8MST)	The 8-minute test is fast, executable, inexpensive, and valid for technical skills assessment
2016	Toprak., *et al*.	Prospective observational	GC and orthopedic residents and assistants of these specialties	Surgical Procedure Feedback Rubric (SPR) - The SPR measures 3 factors: Operating Room Preparation, Technical Skill, and Intrinsic Competencies	Check the effectiveness of the PRS as a tool for evaluating the resident in the intra-op	Validation of the SPR method, mainly to distinguish the levels of knowledge among those evaluated.
2013	Wade., *et al*.	Prospective observational	Residents at the beginning of the first year and at the end of the first year	Suture analysis (filmed procedure) Through the institution's instrument, (technique, time, likert scale, written feedback)	Assess the resident's technical competence	Older residents scored higher on the assessment
2013	Glarner., *et al*.	Prospective observational	Residents (r3 and r5)	Procedure-specific evaluation (key components) + NOTSS (NOTECHS for surgeons) + OSATS during a 2-month rotation and comparative evaluation between the beginning and end of the Caster	Create an instrument that evaluates the resident in technical and non--technical questions	Higher score in relation to all items evaluated for R5 and improvement of the procedure-specific technique (colectomy), without improvement of general skills or NOTECHS.
2012	Balayla., *et al*.	Prospective observational	Interns, residents and assistants	Technical evaluation (check list of the steps of the surgeries: inguinal hernia, CCC VLP, colectomy D)	Observe and compare the less with more experienced and the evolution during the program	Score increases with surgeon experience and variability (standard deviation) within training levels decreases as experience Surgical increased

Appindex 2

Bibliography

1. J Frank., *et al*. "Competency-based medical education: theory to practice". *Medical Teacher* 32.8. (2010): 638-645.

2. J Shalhoub., *et al*. "What evidence is there for the use of workplace-based assessment in surgical training?" *Journal of Surgical Education* 71.6. (2014): 906-915.

3. J Wagner., *et al*. "Use of entrustable professional activities in the assessment of surgical resident competency". *JAMA Surgery* 153.4. (2018): 335-343.

4. J Weller., *et al*. "Systematic review and narrative synthesis of competency- based medical education in anaesthesia". *British Journal of Anaesthesia* 124.6. (2020): 748-760.

5. A Ponton-Carss., *et al*. "Assessment of technical and nontechnical skills in surgical residents". *The American Journal of Surgery* 212.5. (2016): 1011-1019.

6. Gaunt A., *et al*. "Exploring the Role of Self-Motives in Postgraduate Trainees' Feedback-Seeking Behavior in the Clinical Workplace: A Multicenter Study of Workplace-Based Assessments From the United Kingdom". *Academic Medicine* 93.10 (2018): 1576-1583.

7. Shackelford S and Bowyer M. "Modern Metrics for Evaluating Surgical Technical Skills". *Current Surgery Reports* 5 (2017) 24.

8. Barreira MA., *et al*. "Evaluation tool for a gastroenterostomy simulated training1". *Acta Cirúrgica Brasileira* 34.3 (2019): e201900308.

9. Tricco AC., *et al*. "PRISMA extension for scoping reviews (PRISMA- ScR): checklist and explanation". *Annals of Internal Medicine* 169.7 (2018): 467-473.

10. McGowan J., *et al*. "PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement". *Journal of Clinical Epidemiology* 75 (2016): 40-46.

11. Mourad Ouzzani., *et al*. "Rayyan-a web and mobile app for systematic reviews". *Systematic Reviews* 5 (2016): 210.

12. Niitsu H., *et al*. "Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room". *Surgery Today* 43.3 (2013): 271-275.

13. Gofton WT., *et al*. "The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence". *Academic Medicine* 87.10 (2012): 1401-1407.

14. Chevallay M., *et al*. "Implementation and validation of a competency assessment tool for laparoscopic cholecystectomy". *Surgical Endoscopy* 36.11 (2022): 8261-8269.

15. Larson JL., *et al*. "Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents". *Surgery* 138.4 (2005): 640-647.

16. Williams RG., *et al*. "A controlled study to determine measurement conditions necessary for a reliable and valid operative performance assessment: a controlled prospective observational study". *Annals of Surgery* 256.1 (2012): 177-187.

17. C-SATS virtual platform (2023).

18. Cox ML., *et al*. "Validation of the Omni: A Novel, Multimodality, and Longitudinal Surgical Skills Assessment". *Journal of Surgical Education* 75.6 (2018): e218-e228.

19. Vassiliou MC., *et al*. "A global assessment tool for evaluation of intraoperative laparoscopic skills". *The American Journal of Surgery* 190.1 (2005): 107-113.

20. Harriman D., *et al*. "The Resident Report Card: A Tool for Operative Feedback and Evaluation of Technical Skills". *Journal of Surgical Research* 239 (2019): 261-268.

21. Bohnen JD., *et al*. Procedural Learning and Safety Collaborative (PLSC). "The Feasibility of Real-Time Intraoperative Performance Assessment With SIMPL (System for Improving and Measuring Procedural Learning): Early Experience From a Multi-institutional Trial". *Journal of Surgical Education* 73.6 (2016): e118-e130.

22. Wilson AB., *et al*. "Clinical Assessment and Management Examination--Outpatient (CAMEO): its validity and use in a surgical milestones paradigm". *Journal of Surgical Education* 72.1 (2015): 33-40.

23. Steiman J., *et al*. "Measuring Competence in Surgical Training through Assessment of Surgical Entrustable Professional Activities". *Journal of Surgical Education* 75.6 (2018): 1452-1462.

24. Yule S., *et al*. "Development of a rating system for surgeons' non-technical skills". *Medical Education* 40.11 (2006): 1098-1104.

25. Glarner CE., *et al*. "Utilizing a novel tool for the comprehensive assessment of resident operative performance". *Journal of Surgical Education* 70.6 (2013): 813-820.

26. Jung JJ., *et al*. "Nontechnical Skill Assessment of the Collective Surgical Team Using the Non-Technical Skills for Surgeons (NOTSS) System". *Annals of Surgery* 272.6 (2020): 1158-1163.

27. Toprak A., *et al*. "Validation of a novel intraoperative assessment tool: The Surgical Procedure Feedback Rubric". *The American Journal of Surgery* 211.2 (2016): 369-376.

28. IC McManus., *et al*. "Assessment of examiner leniency and stringency ('hawk-dove effect') in the MRCP (UK) clinical examination (PACES) using multi-facet Rasch modelling". *BMC Medical Education* 6 (2006): 42.