



Simulation platforms to assess laparoscopic suturing skills: a scoping review

Elif Bilgic^{1,2} · Motaz Alyafi² · Tomonori Hada² · Tara Landry³ · Gerald M. Fried² · Melina C. Vassiliou²

Received: 29 December 2018 / Accepted: 3 May 2019 / Published online: 14 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Background Laparoscopic suturing (LS) has become a common technique used in a variety of advanced laparoscopic procedures. However, LS is a challenging skill to master, and many trainees may not be competent in performing LS at the end of their training. The purpose of this review is to identify simulation platforms available for assessment of LS skills, and determine the characteristics of the platforms and the LS skills that are targeted.

Methods A scoping review was conducted between January 1997 and October 2018 for full-text articles. The search was done in various databases. Only articles written in English or French were included. Additional studies were identified through reference lists. The search terms included “laparoscopic suturing” and “clinical competence.”

Results Sixty-two studies were selected. The majority of the simulation platforms were box trainers with inanimate tissue, and targeted basic suturing and intracorporeal knot-tying techniques. Most of the validation came from internal structure (rater reliability) and relationship to other variables (compare training levels/case experience, and various metrics). Consequences were not addressed in any of the studies.

Conclusion We identified many types of simulation platforms that were used for assessing LS skills, with most being for assessment of basic skills. Platforms assessing the competence of trainees for advanced LS skills were limited. Therefore, future research should focus on development of LS tasks that better reflect the needs of the trainees.

Keywords Laparoscopy · Suturing · Simulation · Assessment

Laparoscopic suturing (LS) is a skill needed in a variety of advanced laparoscopic procedures that have become a routine part of practice, where LS is used for bowel anastomosis, closure of hiatal defects, handling complications, and other procedures. This skill has been identified as one of the more challenging skills for surgical trainees to master [1, 2]. A needs assessment conducted by Nepomnayshy et al. and Enani et al. identified a gap between LS skills needed in the operating room (OR) and LS skills targeted by various simulation platforms [3, 4]. The main gaps were identified

for suturing under tension, suturing in tighter spaces, performing bowel anastomosis, backhand suturing, and suturing using automated devices. These findings were in conjunction with a survey by Mattar et al. where they found that more than half of the program directors (PD) did not think graduated residents had enough skills to perform LS in the OR [5]. Together these findings all suggest that there is a need to improve the training of the residents when it comes to LS. In order to track trainee progress and make sure that they are competent in performing LS, we need to be able to assess their skills. The purpose of our scoping review is to identify simulation platforms available for assessment of LS skills, and determine the characteristics of the platforms and the LS skills that are targeted.

✉ Melina C. Vassiliou
melina.vassiliou@mcgill.ca

¹ Steinberg Centre for Simulation and Interactive Learning, McGill University, Montreal, QC, Canada

² Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, McGill University Health Centre, 1650, Cedar Avenue, L9. 313, Montreal, QC H3G 1A4, Canada

³ Montreal General Hospital Medical Library, McGill University Health Centre, Montreal, QC, Canada

Methods

Search strategy

We performed a scoping review between January 1997 and October 2017 for full-text articles. Search strategies were developed with a librarian (T.L.). The search was done in MEDLINE, Embase, CENTRAL, CDSR, and PubMed. MEDLINE search strategy ([Appendix 1](#)) was applied to all databases. The search was repeated in MEDLINE till October 2018 for new publications. Only articles written in English or French were included. Additional studies were identified through reference lists. The search terms included “laparoscopic suturing”, “simulation”, and “clinical competence.”

Study selection

Included studies reported data on development and/or validation of simulation tasks for assessment of LS. Studies were excluded if they (1) only included medical students, (2) were assessment of whole procedures, (3) were part of a program for urology or gynecology or involved LS specific for those specialties, or (4) were educational intervention.

Data extraction

Three independent reviewers (E.B., M.A., T.H.) conducted the screening. Extracted information included type of simulator (box trainer, augmented reality, virtual reality), task (in vivo, ex vivo, inanimate), suturing (intracorporeal (ICK), extracorporeal (ECK), continuous, interrupted, hand-sewn, device-assisted), scoring metrics, and sources of validity.

We considered a study as targeting advanced laparoscopic suturing if the tasks went beyond the FLS intracorporeal knot-tying (ICK) task, where participants perform 1 interrupted suture with ICK, with optimal camera angle and instrument location.

Ethics approval

This is a review article. Since we only extracted published data, and did not contact authors for extra information about their studies and ask for additional data, we did not need an ethics approval or gather written consent.

Validity evidence

Validity refers to the evidence surrounding a simulation task that measures LS skills. There are five sources of validity evaluated: content (can the simulation tasks measure suturing skills), response process (can the scoring be done accurately), internal structure (are the scores consistent), relationship to other variables (do the task scores correlate with other assessments or differentiate between training levels), and consequences (what are the implications of incorporating the task for assessment into the training programs) [6, 7].

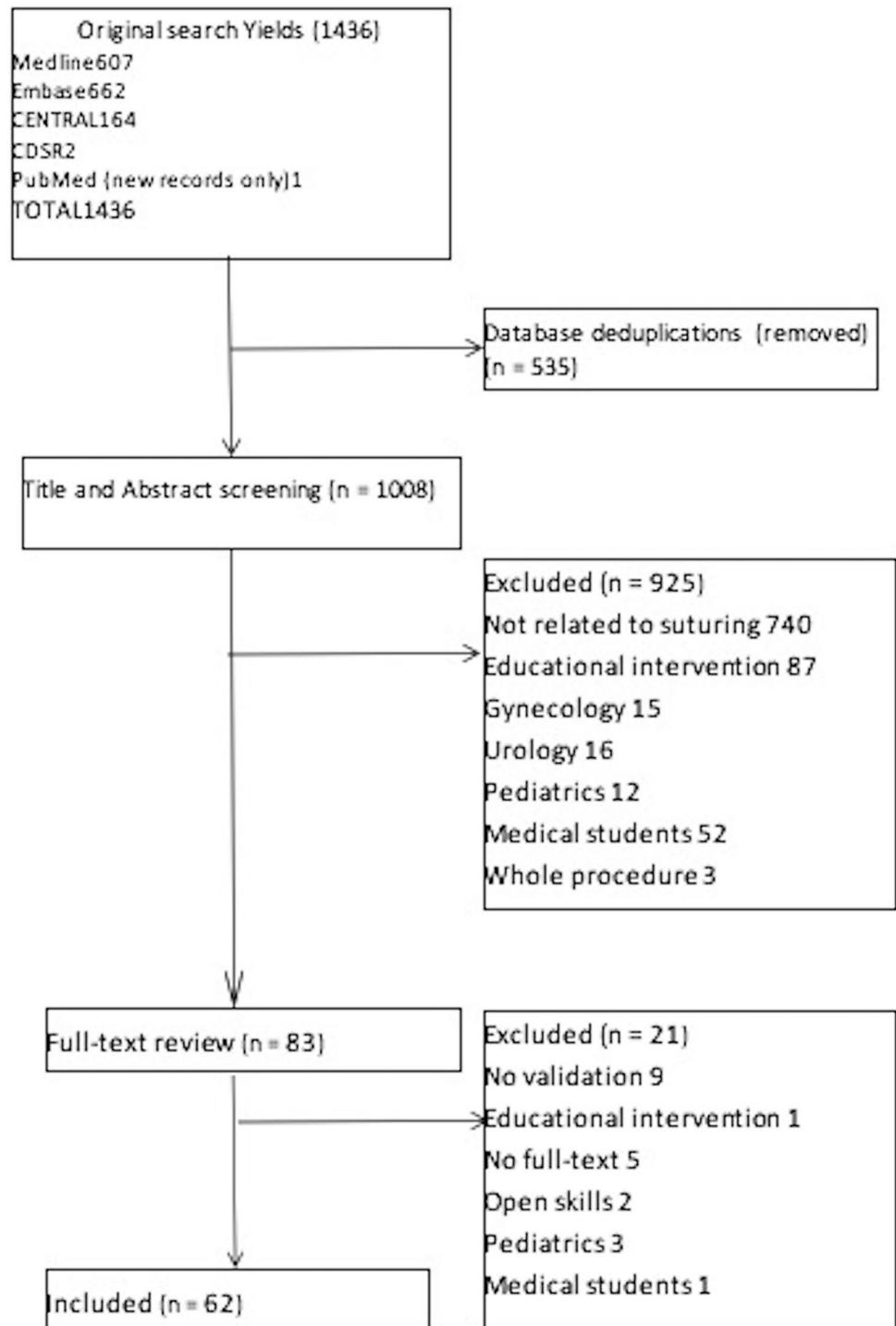
Results

Simulation platforms

Through our search, we included 62 studies for data analysis ([Fig. 1](#)) [8–69]. Some studies used multiple suturing techniques and metrics, which is why the numbers do not add up to 62. Among them, 47 used the box trainer (41 inanimate, 5 ex vivo, 1 not specified (NS)), ten augmented reality, three virtual reality, two in vivo methods, and one cadavers. The majority of the suturing was done using basic ICK techniques (interrupted, 1 suture ICK), similar to the suturing done in the Fundamentals of Laparoscopic Surgery (FLS) ICK task. For knot-tying, 56 studies used intracorporeal (ICK), three extracorporeal (ECK), and 11 no knot-tying. For suturing, seven studies used continuous suturing, 51 interrupted, 57 hand-sewn, and two device-assisted.

Few studies assessed advanced LS skills. Of those studies, 1 assessed continuous suturing with ICK using the Endo Stitch in an inanimate model. One assessed interrupted suturing with ICK in an in vivo porcine model during nissen fundoplication. Four studies used multiple inanimate box trainer tasks to assess various advanced skills (they used a combination of the following tasks); offset angle FLS ICK suture, offset angle backhand ICK suture, confined space FLS ICK suture, needle handling (changing the needle angle), suturing under tension (3 interrupted ICK), and continuous suturing (anastomosis). Two assessed anastomosis techniques with continuous suturing using a porcine intestine model, 1 assessed continuous suturing using a synthetic suturing pad, 1 assessed the anastomosis technique with interrupted suturing using a synthetic intestine model, and 1 assessed ICK in a difficult location on a lamb liver (deep suturing). There was also one major vessel injury (MVI) model where trainees

Fig. 1 Flow chart



had to perform suturing of any kind to stop the bleeding of a synthetic tissue. In terms of metrics, 32 studies used multiple metrics, 0 used time alone, seven time and error alone, 18 motion analysis alone, and four assessment tools alone. More details can be found in Table 1.

Validity evidence

The majority of the evidence came from the internal structure where they evaluated rater consistency and the relationship to other variables where they evaluated differences between different training levels (Table 2). Ten studies

Table 1 Studies reporting data on development and/or validation of simulation tasks for laparoscopic suturing (stars in the first column represent the number of validity sources addressed (1–5 stars); details can be found in Table 2)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Anton N/2018**	In vivo	Porcine Gastrogastric ICK (interrupted, but number NS)	(1) GOALS (2) ISAT (3) Time and error ('objective score')	+/	NS	GOALS and ISATS: (1) could accurately measure performance compared to the time+error, (2) are highly correlated with each other and time+error, (3) high observer agreement. For ISAT, factor analysis identified four unique factors being assessed	Performances were assessed using videos.
Bilgic E/2018**	Box trainer	Sponge ICK (continuous suturing)	(1) Time and error (2) GOALS (3) Endo Stitch assessment tool (EAT)	/+	+	The metrics can distinguish between different levels of experience. High IRR and good correlation between metrics. Experts agree on the value of the task. Expert benchmarks were developed.	Used the Endo Stitch device
Poursartip B/2018*	Box trainer	Synthetic suturing pad (1) Pass needle through tissue (2) ICK (1 suture)	(1, 2) Motion analysis (Potential energy, kinetic energy, work, and combination)	+/	NS	The metrics can distinguish between different levels of experience.	Participants repeated both tasks four times. SIMIS system was used to measure performance.
Sawyer JM/2018*	In vivo (live porcine)	NS ICK (number NS)	(1) GOALS (2) Time and error	+/	NS	Moderate IRR (for full and part videos, low for 30 s).	Suturing videos (full-length, half of video, 30 s video) were assessed by experts and novices. Time and error were assessed live, and GOALS through videos. They compared ratings for different video lengths.

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Takayasu K/2018*	Box trainer	Penrose ICK (1 suture: one double knot, then five single knots)	(1) Motion analysis (upper body motion including joint angle etc.) (2) Time	+/ NS	NS	The metrics can distinguish between experts and novices.	
Wee J/2018*	Box trainer	Penrose FLS ICK (1 suture)	(1) Motion analysis (force metrics) (2) Time and error	+/ NS	NS	Only 1st puncture force metric can distinguish between experts and novices. No significant correlation between force and time and error.	The needle holder is sensorized to measure five forces per suture: two puncture forces, and three knot tightening forces.
Yamaguchi T/2018*	Box trainer	Synthetic suturing pad ICK (number NS)	Motion analysis (time, average velocity, work density, instrument cross time)	+/ NS	NS	The metrics can distinguish between experts and novices.	
Bilgic E/2017*	Box trainer	(1) Transparent circle with six holes (2, 3, 4) Penrose (5, 6) Fabric (3) Offset angle backhand suture (1 suture) (4) Confined space FLS suture (1 suture) (5) Suturing under tension (3 sutures) (6) Continuous suturing	(1) Time and error	+/ +	+	Expert time and error benchmarks were developed. High IRR.	Time and error were assessed live. Experts completed the tasks twice, and mean time and median error scores were used. (Same tasks as Watanabe Y/2015 and three from Nepomnayshy D/2015)
Leeds S/2017*	Box trainer	Synthetic suturing pad ICK (continuous suturing)	(1) Time (2) Dots on target (DoT) (3) Total deviation (D) (4) Number of attempts to reach proficiency	+/ +	+	Experts had significantly better time and took less attempts to reach proficiency. DoF and D was better for experts, but not significantly.	Participants completed the task using Endo360 and traditional laparoscopic technique. The results apply to both devices.

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Kowalewski KF/2017**	Box trainer	Silicone suture pad ICK (4 sutures)	(1) Motion analysis (author's new system called iSurgeon, includes multiple parameters) (2) OSATS (Chang et al., generic and LS specific items)	+/ +	+	iSurgeon system and OSATS can distinguish between experts and novices. The two metrics had strong correlation.	
McKendy K/2017**	Box trainer	(1) Transparent circle with 6 holes (2, 3) Fabric (1) Needle handling (no knot-tying) (2) Suturing under tension (3 sutures) (3) Continuous suturing	(1) Time and error (2) FFT	+/ +	NS	Metrics were developed in an iterative fashion, and can distinguish between novice and experts. High IRR, internal consistency, and correlation between metrics.	Time and error were done live, and FFT using videos. The results are true for all three tasks (Same tasks as Watanabe Y/2015)
Rosser JC/2017*	Box trainer	Porcine intestine ICK (number of sutures NS)	Time and error	+/ +	NS	ICK task correlates with case experience and Monkey Ball videogame	They have two video games (Monkey Ball and Underground), FLS peg transfer, pea drop, and ICK. Result for each task is separate.
Sleiman Z/2017**	(1) Box trainer (home-made) (2) Box trainer (stand-ard)	(1, 2) Synthetic suturing pad ICK (1 suture)	(1, 2) Stitch out of the dot, tear in the tissue	+/ +	+	The metrics can distinguish between experts and novices. Participants found value in homemade trainer for home training, and value in both trainers for training in general	They have four tasks; one being ICK. Result for each task is separate
Yeung C/2017*	Box trainer	Penrose ICK (1 suture)	(1) Overall performance (Likert scale 1–5) (2) Number of times raters changed the score	+/ +	NS	Consistency between expert and non-expert raters was high. Experts changed their scores fewer times (not reach statistical significance)	

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Deal SB/2016*	Box trainer	FLS ICK (1 suture)	GOALS 20pt	+/	NS	Assessment using crowd-source has some reliability	
Kowalewski KF/2016**	Box trainer	ICK (4 sutures)	(1) Motion analysis (author's new system called iSurgeon) (2) LSIKT-GRS+LSIKT-CL	+/	+	iSurgeon system has some validity to assess skill	
Sanchez-Margallo JA/2016**	Box trainer	ICK (3 sutures)	(1) Moorthy CL 27pt (2) Motion analysis (Micron Tracker)	+/	NS	Validity evidence was gathered for assessment using motion metrics, but need methods to assess quality of the suturing.	
Poursartip B/2016*	Box trainer	pass needle through incision ICK (1 suture)	Motion analysis (SIMIS system); metrics are potential energy, kinetic energy, work)	+/	NS	Trainees with more laparoscopic case experience scored better	Participants performed the tasks four times. Their 3rd trial scores were used
Zdichavsky M/2016*	Box trainer	ICK (continuous suturing)	(1) Time (2) Accuracy (3) Anastomatic width (4) Pressure resistance	+/	+	Trainees with more laparoscopic case experience scored better	The needle driver is a new steerable instrument called r2. It has active tip deflection, and tip and shaft rotation.
Veneziano D/2016**	Box trainer	(1) ICK (1 suture) (2) They could use any technique to stop the bleeding	(1) Time and error (FLS scoring) (2) Blood loss, number of stitches	+/	+	Trainees in higher training levels scored better (statistical significance not reported). Trainees agreed that MVI model can assess repair of MVI skill, is realistic, and should be included in the curriculum	The MVI model was perfused with synthetic blood. The room had low lights and pulse sounds to simulate OR.

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Buckley CE/2015*	Box trainer (ProMIS augmented reality)	Synthetic suturing skin ICK (1 suture)	(1) Computer-generated metrics in the simulator (2) OSATS 25pt (3) Time and error (FLS scoring) (4) Zone calculations from a new software	+/ +	+	The zone calculation is a proper metric for this task	
Chang OH/2015**	Box trainer	Penrose FLS ICK (1 suture)	(1) LSIKT-GRS (2) LSIKT-CL	+/ NS	NS	The assessment tools are able to measure suturing skills	
Nepomnyashy D/2015*	Box trainer	FLS tasks: (1) FLS peg transfer (2) FLS pattern cut (3) FLS ICK ALS tasks: (4) Offset simple suture (5) Offset backhand suture (6) Offset peg transfer (7) 180 degree 'reversely' peg transfer (8) Confined suture (9) Peg transfer on the block (10) Circle cut on the block	Time and error for all	+/ NS	NS	The metrics can distinguish between experts and novices. Experts agree on the value of the tasks.	Separate and Compound FLS and ALS scores were generated (summing task scores for the two). We are focusing on the suturing tasks for results
Stefanidis D/2015*	Box trainer	Penrose FLS ICK (1 suture)	Time and error (FLS scoring)	+/ +	+	PGY level and case experience were good predictors of simulation performance.	They have three laparoscopic (1 is FLS ICK) and five open tasks. Result for each task is separate
Trudeau MO/2015*	Box trainer	Penrose ICK (1 suture)	Motion analysis (velocity, acceleration, range)	+/ NS	NS	Trainees with more laparoscopic case experience scored better (range was not statistically different)	

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Uemura M/2015*	Box trainer (Augmented reality)	Synthetic intestinal model ICK (3 sutures)	(1) Time (2) Air pressure leak (3) Number of full-thickness sutures (4) Suture tension (5) Wound-opening area	+/ +	+	The metrics can distinguish between experts and novices. Expert benchmarks were developed	The simulation platform is called Suture Simulator Instruction Evaluation Unit.
Watanabe Y/2015**	Box trainer	(1) Transparent circle with 6 holes (2, 3) Fabric (1) Needle handling (no knot-tying) (2) Suturing under tension (3 sutures) (3) Continuous suturing	(1) Time and error	+/ NS	NS	Metrics can distinguish between different levels and case experiences. High IRR and internal consistency. Experts agree on the value of the tasks.	
Kowalewski TM/2014**	Box trainer Simulab EDGE platform	Penrose FLS ICK (1 suture)	(1) Time and error (2) p-OSATS (3) Motion analysis (automatically calculated by the EDGE platform)	+/ NS	NS	The various metrics used had good correlation with each other, and higher training levels/case experience scored better, and p-OSATS showed good interrater reliability.	
Trejos AL/2014*	Box trainer	Foam and silicone Needle driving (no knot-tying) ICK (1 suture)	Motion analysis (SIMIS system; metrics based on time, position, and force)	+/ NS	NS	Force-based metrics are better able to differentiate experts from novices	The task is a procedure that has five subtasks that are assessed separately; 1 driving needle through tissue, 1 ICK. Result for each task is separate. Participants completed the procedure 4 times
Stefaniadis D/2014*	Box trainer	Penrose FLS ICK (1 suture)	Time and error (FLS scoring)	+/ +	+	Validity evidence was gathered for the task	They have three laparoscopic(1 is FLS ICK) and five open tasks. Result for each task is separate. Participants completed each task two times

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Lusch A/2014*	Box trainer	(1) Rings through ring without knot-tying (2) ICK (1 suture) (3) ICK (multiple sutures)	Quantity score: (1) Number of passed rings (2) Number of ties within knot (3) Number of suture throws Quality score: (1) Number of missed attempts (2, 3) Distance between knot and suture	+/ NS	NS	Validity evidence was gathered for the 2D and 3D systems	The study was comparing 2D vs 3D systems. So the validity evidence is for both. They have six tasks, three suturing related. Participants completed all tasks in 2D and 3D. Result for each task and system is separate.
Escamirosa F/2014***	Box trainer (EndoVis augmented reality training system)	ICK (1 suture)	Metrics of the simulator	+/ NS	NS	Validity evidence was gathered for the task	They have four tasks; 1 being ICK. Result for each task is separate
Romero P/2014*	Box trainer	Synthetic suturing pad ICK (3 sutures)	(1) Time (2) Knot quality (5pt scale) (3) Accuracy (4) Moorthy CL 23 pt	+/ +	+	Experts had better scores than junior trainees	
Egi H/2013**	Box trainer	ICK (1 suture)	(1) Moorthy CL 29pt (2) OSATS Buckley 2015	+/ NS	NS	Good correlation between the LS task scores and HUESAD(this is an augmented reality simulator & doesn't assess LS) motion analysis task scores	
Boza C/2013*	Box trainer	Porcine intestine (Jejuno-jejunostomy) ICK (continuous suturing)	(1) Motion analysis (ICSAD (time, path length, total no of movements)) (2) ALRYGB	+/ NS	NS	Strong correlation between performance in the simulator and the OR (completed jejunostomy of a LRYGB and assessments were identical to simulator)	

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Hennessey I/2013**	(1) Box trainer(FLS) (2) Box trainer (eoSim)	(1) Penrose (2) NS(looks like a fabric)	(1) Time and error (FLS scoring) (2) NS	+/ -	+	Validity evidence was gathered for the two platforms	There were 3 tasks: object transfer, cutting, ICK. Result for each task is separate. Participants completed both platforms, but the order of which one they start with was randomly selected.
Bahsoun AN/2013*	Box trainer	Hoops	Passing needle through 3 hoops without knot-tying	-	NS	Experts found that the trainer had high training capacity and performance (video, light etc.)	The box is made of cardboard, and use iPad2 as camera and monitor.
Palter V/2012*	VR(LAPSIM)	ICK (1 suture)	Metrics of the simulator(time, path length, angular path)	+/ -	+	Expert benchmarks for the tasks were developed	First, they did a Delphi study to determine which tasks on the VR simulator should be included in the proficiency-based VR technical skills curriculum for colorectal surgery. Then, experts completed the tasks to determine benchmarks (8 tasks are included in the final curriculum, 1 suturing. Result for each task is separate).
Pagador JB/2012*	Box trainer	Carcass stomach	Augmented reality haptic (ARH, motion analysis)	+/ -	NS	Different metrics showed significant difference between the levels for the four subtasks.	LS is decomposed into 4 subtasks; needle puncture, first knot, second knot, third knot.
Sharma M/2012**	Fresh Frozen Cadaver (FFC)	Mesenteric rifts	GOALS 20pt	+/ -	+	The metrics can distinguish between experts and novices. High inter-rater reliability for ICK task. Expert benchmarks were developed	They have five tasks, including ICK and ECK. Result for each task is separate.

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Strickland A/2011*	Box trainer (ProMIS augmented reality)	Lamb liver ICK (2 sutures (one at an easier and one at a more difficult location))	Metrics of the simulator (time, path length)	+/ +	+ +	The metrics can distinguish between experts and novices	They have four tasks; 1 easy stitch, 1 hard stitch. Results for each task is separate
Hiemstra E/2011*	Box trainer	Synthetic suturing pad ICK (1 suture, they performed the task three times consecutively)	Motion analysis (TrEndo; time, path length, motion in depth, motion smoothness)	+/ +	+ +	Experts had better scores than residents and medical students.	LS skills improved from 1st to 3rd trial for medical students and residents, but experts were consistent.
Kobayashi SA/2011**	Box trainer	Penrose FLS ICK (1 suture) FLS ECK (1 suture)	Time and error (FLS scoring)	+/ +	+ +	The metrics can distinguish between experts and novices. Experts found the simulator realistic and could help improve skill.	They have 5 tasks, including ICK and ECK. Result for each task is separate
Zheng B/2010*	Box trainer	Synthetic soft tissue ICK (interrupted, as many sutures as possible in 6 min) while also responding to a visual cue	(1) Time and error (FLS scoring) (2) Number of sutures completed (3) Error scoring for the visual detection task (secondary)	+/ NS	NS	Experienced surgeons performed more sutures, had higher quality, and scored better in the secondary task.	While performing the suturing tasks, participants were asked to respond to visual cues correctly, and authors hypothesize that experts will have more space in their cognitive capacity to perform this secondary task.
Yamaguchi S/2010*	Box trainer	Synthetic suturing pad ICK (1 suture)	Motion analysis (AUORA; time, path length, average speed)	+/ +	+ +	Experts had better scores than residents	
Horeman T/2010*	Box trainer	Synthetic skin model Needle driving (no knot-tying required)	Motion analysis (force)	+/ NS	NS	The force metric can distinguish between experts and novices	Participants completed the task twice.
Botden SM/2009**	Box trainer (ProMIS augmented reality)	Synthetic suturing pad ICK (1 suture)	(1) Time (2) Knot strength (3) Time spent in correct area (4) Botden 2009 LS specific tool	+/ NS	NS	Using augmented reality is a way to measure LS skills	

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Kroeze SGC/2009**	Box trainer	Synthetic suturing pad ICK (1 suture)	Moorthy CL 29pt	+/ +	+	There was a relationship between score and PGY level. Validity evidence was gathered for the 2 platforms	
Lin D/2009**	(1) Box trainer(ProMIS augmented reality) (2) VR simulator (SurgicalSIM)	(1, 2) ICK (1 suture)	Metrics of the simulators	+/ +	+		
Xeroulis G/2009*	Box trainer	Penrose FLS ICK (1 suture)	(1) Time and error (FLS scoring) (2) Motion analysis (ICSAD)	+/ +	+	The metrics can distinguish between experts and novices, and they correlate significantly.	They have four tasks; 1 being ICK. Result for each task is separate
Dayan AB/2008**	Box trainer	Synthetic suturing pad ICK (1 suture)	(1) Moorthy CL 27pt (2) Time	+/ +	+	The simple, low-cost laparoscopic training platform has good validity evidence	There were 3 tasks: rope passing, peg transfer, and intracorporeal knot-tying. Participants completed all 3 tasks, but results separate for each. Short warm-up session. Max time 15 min for LS
Oostema JA/2008*	Box trainer (ProMIS augmented reality)	NS ICK (1 suture)	Metrics of the simulator (time, path length, smoothness)	+/ +	+	Participants with more laparoscopic case experience scored better	They have 4 tasks; 1 being ICK. They performed 3 repetitions. Result for each task is separate
Zheng B/2007**	Box trainer	Penrose ICK (1 suture)	(1) Time and accuracy (LISETT score) (2) Self-rated team quality scores	+/ +	NS	Validity evidence was gathered for the LISETT system that is meant to enhance team skills.	They have 2 tasks; peg transfer and ICK. The results are for the combination score of the 2 tasks
Stefanidis D/2007*	Box trainer	Penrose ICK (interrupted, as many sutures as possible in 10 min) while also responding to a visual cue	(1) Time and error (FLS scoring) (2) Number of sutures completed (3) Error scoring for the secondary task	+/ +	NS	Experienced surgeons performed more sutures, had higher quality, and scored better in the secondary task.	While performing the suturing tasks, participants were also asked to complete a visual-spacial secondary task for attention (2 tasks performed simultaneously).

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Bolden SM/2007**	(1) Box trainer (ProMIS augmented reality) (2) VR(LAPSIM)	(1) Synthetic suturing pad (2) –	Metrics of the simulators	+/ –	+	Participants found augmented reality (AR) to be more realistic and had better training quality. 4/5 AR metrics and 1/5 VR metrics can distinguish experts and novices.	Participants performed 2 tasks (1 LS) in each simulator. Result for each task and simulator is separate
Broe D/2006**	Box trainer (ProMIS augmented reality)	Passing needle through 4 hoops without knot-tying	OSATS 25pt	–	+	Senior residents scored better than junior, and there was good inter-rater reliability.	There were 3 tasks: Laparoscopic orientation, dissection and suturing. The results are for the combination of 3. They also had a CL for each of the 3 tasks, however, we do not know what the items are, so we only looked at GRS
Dubrowski A/2006*	Box trainer	(No knot-tying, only passing needle from left part of the tissue to right)	Motion analysis	–	+	Most motion analysis metrics were sensitive to residency level	They performed suturing 10 times, without knot-tying
Van Sickle KR/2005*	Box trainer (ProMIS augmented reality)	(Passing needle through 5 paired circles without knot-tying)	Metrics of the simulator (time, path length, smoothness)	–	+	The metrics can distinguish between experts and novices	
Duffy AJ/2005*	VR(LAPSIM)	ICK (1 suture)	Metrics of the simulator	+/ –	+	The VR metrics can distinguish between experts and novices	They have 8 tasks, 1 being ICK. Result for each task is separate
Moorthy K/2004*	Box trainer	ICK (2–3 sutures)	(1) Moorthy CL 29pt (2) Motion analysis (ICSAD (time, path length))	+/ –	+	The various metrics used had good correlation with each other	They had experts and novices. Novices received video-based instructions before completing the task.

Table 1 (continued)

Author/Year	Platform	Tasks	Metrics	H/D	Warm-up/teaching	Results	Extra notes
Figert PL/2001*	Box trainer	Organ-shaped foam rubber ICK (3 sutures) for 3 different knot techniques	(1) Time (2) 14 item error rating-tool(LS-ERS)	+/ +	+	Trainees with more laparoscopic case experience scored better	They completed a pre-assessment session (1) Didactic session(instruments, trocar and camera placement, knot-tying techniques) (2) Demonstration & written instructions on the techniques
Keyser EJ/2000*	Box trainer (2 different boxes were used)	Penrose FLS ICK (1 suture) FLS ECK (1 suture)	Time and error (FLS scoring)	+/ +	+	Trainees with more laparoscopic case experience scored better in both boxes, and there was moderate correlation between the 2 box scores	They have 7 tasks, including ICK and ECK. Result for each task is separate

H/D Hand-sewn/Device-assisted, *LS* Laparoscopic suturing, *ICK* Intracorporeal knot-tying, *FLS* Fundamentals of Laparoscopic Surgery, *GOALS* Global Operative Assessment of Laparoscopic Surgery, *ISAT* Intraoperative Suturing Assessment Tool, *CL* Checklist, *OSATS* Objective Structured Assessment of Technical Skills, *ERS* Error Rating Scale, *MS* Not Specified, *p-OSATS* Psychomotor-OSATS, *ICSAD* Imperial College Surgical Assessment Device, *ALRYGB* Assessment of Laparoscopic Roux en-Y Gastric Bypass, *GRS* Global Rating Scale, *LSIKT* Laparoscopic suturing and intracorporeal knot-tying, *LSETT* Legacy Inanimate System for Endoscopic Team Training, *VR* Virtual Reality, *SIMIS* Sensorized instrument-based minimally invasive surgery system, *pt Point*, *ECK* Extracorporeal knot-tying, *MVI* Major vessel injury, *OR* Operating room, *IRR* Inter-rater reliability, *SIMIS* Sensorized Instrument-based Minimally Invasive Surgery, *ALS* Advanced Laparoscopic Surgery, *FFT* Formative Feedback Tool

Table 2 Validity evidence of studies that focus on validation of the laparoscopic suturing assessment platforms

Author/Year	Sources of validity targeted
Anton N/2018	C (cognitive task analysis, expert opinion) RP (rater training, clarification of metrics) I (IRR for assessment tools, factor analysis for ISAT) R (correlation between the metrics, discriminatory accuracy of metrics)
Bilgic E/2018	C (expert opinion, literature review) I (IRR) R (case experience; correlation of the metrics, expert benchmarks)
Poursartip B/2018	C (literature review) R (training level)
Sawyer JM/2018	RP (rater training, lowering rater bias) I (IRR)
Takayasu K/2018	R (training level)
Wee J/2018	R (case experience)
Yamaguchi T/2018	RP (iterative metric development) R (training level)
Bilgic E/2017 (Same tasks as Watanabe Y/2015)	C (reference to previous study) RP (expert benchmark) I (IRR)
Leeds S/2017	R (training level)
Kowalewski KF/2017	R (training level; correlation between metrics) I (test–retest reliability for both metrics)
McKendy K/2017 (Same tasks as Watanabe Y/2015)	C (reference to previous study, expert opinion) RP (iterative metric and scale development) I (IRR, internal consistency, correlation of the metrics) R (training level)
Rosser JC/2017	R (case experience)
Sleiman Z/2017	C (expert opinion) R (training level)
Yeung C/2017	I (IRR, test–retest)
Deal SB/2016	I (IRR, internal consistency)
Kowalewski KF/2016	R (case experience for both metrics; correlation of the metrics) I (test–retest for both metrics)
Sanchez-Margallo JA/2016	R (Case experience; correlation between the metrics) I (IRR for assessment tool)
Poursartip B/2016	R (case experience)
Zdichavsky M/2016	R (case experience)
Veneziano D/2016	C (participant opinion for MVI model) R (training level)
Buckley CE/2015	R (case experience for zone; correlation of the metrics)
Chang OH/2015	R (training level) I (rater reliability)
Nepomnayshy D/2015	R (training level)
Stefanidis D/2015	R (training level/case experience)
Trudeau MO/2015	R (case experience)
Uemura M/2015	R (training level; expert benchmarks)
Watanabe Y/2015	C (expert opinion) I (IRR, internal consistency) R (training level, case experience)
Kowalewski TM/2014	R (metrics correlated with each other; higher training levels/case experience scored better) I (only for p-OSATS, rater reliability)
Trejos AL/2014	R (training level)
Stefanidis D/2014	R (training level/case experience)
Lusch A/2014	R (training level)

Table 2 (continued)

Author/Year	Sources of validity targeted
Escamirosa F/2014	C (expert opinion from a questionnaire) R (training level) I (Internal consistency for 4 tasks together)
Romero P/2014	R (training level)
Egi H/2013	R (correlation between metrics) I (IRR for the assessment tools)
Boza C/2013	R (correlating simulation and OR scores for ICSAD and ALRYGB)
Hennessey I/2013	C (experts said both simulators are good for training) R (training level; correlation between FLS and eoSim ICK)
Bahsoun AN/2013	C (expert opinion)
Palter V/2012	R (expert benchmark)
Pagador JB/2012	R (training level)
Sharma M/2012	R (training level; expert benchmarks) I (IRR for ICK task)
Strickland A/2011	R (training level)
Hiemstra E/2011	R (training level)
Kobayashi SA/2011	C (expert opinion) R (training level)
Zheng B/2010	R (training level)
Yamaguchi S/2010	R (training level)
Horeman T/2010	R (training level)
Botden SM/2009	R (training level; correlation between the metrics) I (IRR for assessment tool)
Kroeze SGC/2009	R (training level) I (IRR)
Lin D/2009	C (expert opinion) R (training level)
Xeroulis G/2009	R (training level; correlation between the metrics)
Dayan AB/2008	R (case experience for both metrics) I (IRR)
Oostema JA/2008	R (case experience)
Zheng B/2007	C (expert and participant opinion) R (training level/case experience; correlation between LISETT score and team quality)
Stefanidis D/2007	R (training level)
Botden SM/2007	C (participant opinion for AR) R (training level)
Broe D/2006	R (training level) I (IRR)
Dubrowski A/2006	R (training level)
Van Sickle KR/2005	R (training level)
Duffy AJ/2005	R (training level)
Moorthy K/2004	R (case experience; metrics correlated with each other)
Figert PL/2001	R (case experience)
Keyser EJ/2000	R (case experience, correlation between 2 platforms)

C Content, RP Response Process, I Internal structure, R Relationship to Other Variables, CO Consequences, IRR Inter-rater reliability, ISAT Intraoperative Suturing Assessment Tool, p-OSATS Psychomotor-Objective Structured Assessment of Technical Skills, OR Operating room, ICSAD Imperial College Surgical Assessment Device, ALRYGB Assessment of Laparoscopic Roux en-Y Gastric Bypass, ICK Intracorporeal knot-tying, FLS Fundamentals of Laparoscopic Surgery, LISETT Legacy Inanimate System for Endoscopic Team Training

investigated evidence for content through expert opinion, five studies investigated response process, and no study investigated consequences.

Discussion

This review identified studies that developed and/or provided validity evidence for simulation tasks in the context

of assessment of LS skills. The first point that we identified is that most suturing tasks were based on inanimate models, targeting basic LS skills such as 1 interrupted suture with intracorporeal knot-tying. One of the platforms most often used was the FLS ICK task, or studies modifying the ICK task to fit their model. Regarding advanced LS skills, such as suturing under tension, bowel anastomosis, backhand suturing, suturing in tight spaces, and suturing using automated devices, there were various *ex vivo* and inanimate models with varying validity evidence to support their usage for assessment [4]. In addition, none of the studies addressed concerns regarding cost of the simulation platforms.

Metrics are an essential component to simulation since they allow us to objectively assess learner performance and make sure that learners are competent in a given domain, such as LS. Various metrics could have a role in high-stakes assessment (measuring competence, decision-making) or low-stakes assessment (providing feedback), and the metrics have to be linked to the purpose of assessment [70, 71]. Among the studies identified, the majority of the metrics used for assessment were time and error, motion analysis, and assessment tool scores. Time and error metrics are easy to implement and there are plenty of data supporting their use for assessment. But if the assessment's purpose was to provide feedback, time and error metrics only target speed and the end product, which limits learning as the process taken to reach the end product is just as important as the end product itself and evaluating the process discourages formation of bad habits inappropriate to the clinical setting [72, 73]. Motion analysis is a combination of computer-generated metrics. Although motion analysis removes the human factor and improves score consistency, interpreting the meaning behind the scores is not always clear; just because someone had a similar path length to experts does not mean that their end product is clinically sufficient (e.g., knots do not come off, there is no leak, etc.) [74, 75]. Finally, assessment tools can provide meaningful feedback and capture the process of how someone achieved the end product, but they require raters, which could be resource intensive [76]. None of the identified studies used the assessments for high-stakes evaluation of skill level. They all provided evidence in the context of assessing competence without specifying the purpose of assessment. As stated above, having a clear purpose is important when choosing the measures of assessment. Overall, the majority of the time, metrics were a combination of various types. The reason for using different types of metrics could be that due to the complex nature of the LS skill, educators are trying to capture a more complete picture of the trainee performance.

Regarding the validity evidence gathered for LS skills assessment, most studies investigated the internal structure through rater reliability and relationship to other variables by comparing scores of different training groups or correlation

of various metrics, with minimal emphasis of the other 3 sources of validity. Within this validity framework, all sources add something important and the more sources of validity investigated, the more robust the evidence for the tools' potential to measure LS skills. To provide evidence for the content, the tasks could be developed with experts in the field in order to make sure that the skills assessed represent skills needed in the clinical setting. Furthermore, to provide evidence for the response process, steps could be taken to ensure that the scoring is done accurately, through rater training and clarification of what the various performance scores mean. Evidence for the consequences is not addressed much in the literature because it requires a longitudinal investigation of the implications. For example, if the assessor sets pass/fail standards for the assessment, the consequences involve what happens to the trainees who receive a fail grade and what steps are taken to make sure that they can pass. Above all, what we have to understand is that validity is not about the task, it is about the assessment and interpretation of the score trainees receive. This relates to the metrics of the tasks, since metrics construct the score [6]. Through the five sources, educators build an argument about the interpretation of the assessment depending on the purpose of using the simulation task: high-stakes versus low-stakes assessment (validity of the interpretations that we made). This is important since evidence gathered in each of the five sources may differ depending on the purpose. Additionally, in this review, we did not investigate simulations that were used as educational intervention, which might also require different validity evidence. Therefore, establishing a context-dependent validity evidence is important so that educators can accurately reach a conclusion from the evidence gathered. As stated in the previous paragraph, even though all of the studies provided evidence in the context of assessing LS competence, they were not clear on the purpose of assessment.

Among the 62 studies that were analyzed, we identified a lot of different simulation tasks that were used for assessment purposes: box trainer, VR, and augmented reality. In addition, various types of metrics and suturing and knot-tying types were used. The variety of tasks and metrics illustrate a lack of consensus regarding the best way to incorporate simulation platforms to assess LS skills. However, it is also evident from our review that all platforms had varying degrees of validity evidence in the context of assessing LS skills. Therefore, it is more important to consider the limitations/advantages of each platform and choose a platform accordingly. For example, the cost associated with the platforms and the feasibility of obtaining the necessary apparatus may be limiting for some programs. VR is expensive, yet it allows easy assessment of trainees while performing a whole procedure without needing live animal models. Box trainers with *ex vivo* tissue might increase the fidelity of the

task, yet it is hard to preserve the tissues and they are more labor intensive. Box trainers with inanimate models such as penrose and fabrics are a much cheaper platform, which could explain why 73% of the analyzed studies used them, and there is ample evidence to suggest that they are an effective assessment platform.

In summary, there is no one optimal way to assess LS skills, so researchers should develop their own platforms, or chose from one of the existing platforms identified in this review, taking into account their individual context of assessment (type of simulation and metrics, where will the assessment take place, who will be the rater, who will be assessed etc.), purpose of assessment (formative versus summative), and the validity evidence for each of the assessment conditions.

Conclusion

In conclusion, we identified simulation platforms used for assessing LS skills. There were plenty of different platforms, yet platforms that could assess the competence of trainees for advanced LS skills identified by Nepomnayshy et al. (suturing under tension, anastomosis, backhand suturing, suturing using an automated device, suturing in tight spaces) were limited in number and validity evidence. Therefore, future research should focus on assessment of advanced LS skills that better reflect the current practice environment.

Acknowledgements We would like to thank the members of the Henry K.M. DeKuyper Education Centre, Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, and Steinberg Centre for Simulation and Interactive Learning for their help in conducting this study.

Funding There was no funding or financial support for this work.

Compliance with ethical standards

Disclosures The Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation received an unrestricted educational grant from Medtronic Canada. Elif Bilgic, Motaz Alyafi, Tomonori Hada, Tara Landry, Gerald M Fried, and Melina C Vassiliou have no relevant conflicts of interest or financial ties to disclose.

Appendix 1: Medline search strategy

1. Exp Laparoscopy/ (90382)
2. Suture Techniques/ (43044)
3. Sutures/ (16421)
4. 2 or 3 (54604)
5. 1 and 4 (2898)

6. ((Extracorporeal or extra-corporeal or intracorporeal or intra-corporeal or laparo*) adj3 (suture* or knot*)).tw,kf. (2117)
7. 5 or 6 (4242)
8. ed.fs. (261740)
9. Education, medical, graduate/ (26742)
10. “Internship and residency”/ (44310)
11. Exp Teaching/ (80605)
12. Exp Learning/ (362731)
13. (Curricul* or educat* or instruct* or learn* or teach* or train* or tuition).tw,kf. (1305770)
14. or/8-13 (1650389)
15. 7 and 14 (747)
16. Limit 15 to English (696)
17. Remove duplicates from 16 (653)
18. Limit 17 to yr=“2000 Current” (607)

References

1. Palter VN, Orzech N, Aggarwal R, Okrainec A, Grantcharov TP (2010) Resident perceptions of advanced laparoscopic skills training. *Surg Endosc* 24:2830–2834
2. Rattner DW, Apelgren KN, Eubanks WS (2001) The need for training opportunities in advanced laparoscopic surgery. *Surg Endosc* 15:1066–1070
3. Enani G, Watanabe Y, McKendy KM, Bilgic E, Feldman LS, Fried GM, Vassiliou MC (2017) What are the training gaps for acquiring laparoscopic suturing skills? *J Surg Educ* 74:656–662
4. Nepomnayshy D, Alseidi AA, Fitzgibbons SC, Stefanidis D (2016) Identifying the need for and content of an advanced laparoscopic skills curriculum: results of a national survey. *Am J Surg* 211:421–425
5. Mattar SG, Alseidi AA, Jones DB, Jeyarajah DR, Swanson LL, Aye RW, Wexner SD, Martinez JM, Ross SB, Awad MM, Franklin ME, Arregui ME, Schirmer BD, Minter RM (2013) General surgery residency inadequately prepares trainees for fellowship: results of a survey of fellowship program directors. *Ann Surg* 258:440–449
6. Ghaderi I, Manji F, Park YS, Juul D, Ott M, Harris I, Farrell TM (2015) Technical skills assessment toolbox: a review using the unitary framework of validity. *Ann Surg* 261:251–262
7. Downing S (2003) Validity: on meaningful interpretation of assessment data. *Med Educ* 37:830–837
8. Leeds SG, Wooley L, Sankaranarayanan G, Daoud Y, Fleshman J, Chauhan S (2017) Learning curve associated with an automated laparoscopic suturing device compared with laparoscopic suturing. *Surg Innovat* 24:109–114
9. Sánchez-Margallo JA, Sánchez-Margallo FM, Oropesa I, Enciso S, Gómez EJ (2016) Objective assessment based on motion-related metrics and technical performance in laparoscopic suturing. *Int J Comput Assist Radiol Surg* 12:1–8
10. Kowalewski KF, Hendrie JD, Schmidt MW, Garrow CR, Bruckner T, Proctor T, Paul S, Adigüzel D, Bodenstedt S, Erben A, Kennigott H, Erben Y, Speidel S, Müller-Stich BP, Nickel F (2016) Development and validation of a sensor- and expert model-based training system for laparoscopic surgery: the iSurgeon. *Surg Endosc* 31:1–11
11. Deal SB, Lendvay TS, Haque MI, Brand T, Comstock B, Warren J, Alseidi A (2016) Crowd-sourced assessment of technical skills:

- an opportunity for improvement in the assessment of laparoscopic surgical skills. *Am J Surg* 211:398–404
12. Chang OH, King LP, Modest AM, Hur HC (2016) Developing an objective structured assessment of technical skills for laparoscopic suturing and intracorporeal knot tying. *J Surg Educ* 73:258–263
 13. Stefanidis D, Gardner C, Paige JT, Korndorffer JR Jr, Nepomnyashy D, Chapman D (2015) Multicenter longitudinal assessment of resident technical skills. *Am J Surg* 209:120–125
 14. Escamirosa FP, Flores RM, Garcia IO, Vidal CR, Martinez AM (2015) Face, content, and construct validity of the EndoViS training system for objective assessment of psychomotor skills of laparoscopic surgeons. *Surg Endosc* 29:3392–3403
 15. Buckley CE, Kavanagh DO, Nugent E, Ryan D, Traynor OJ, Neary PC (2015) Zone calculation as a tool for assessing performance outcome in laparoscopic suturing. *Surg Endosc* 29:1553–1559
 16. Trejos AL, Patel RV, Malthaner RA, Schlachta CM (2014) Development of force-based metrics for skills assessment in minimally invasive surgery. *Surg Endosc* 28:2106–2119
 17. Stefanidis D, Grewal H, Paige JT, Korndorffer JR, Scott DJ, Nepomnyashy D, Edelman DA, Sievers C (2014) Establishing technical performance norms for general surgery residents. *Surg Endosc* 28:3179–3185
 18. Romero P, Brands O, Nickel F, Muller B, Gunther P, Holland-Cunz S (2014) Intracorporeal suturing—driving license necessary? *J Pediatr Surg* 49:1138–1141
 19. Lusch A, Bucur PL, Menhadji AD, Okhunov Z, Liss MA, Perez-Lanzac A, McDougall EM, Landman J (2014) Evaluation of the impact of three-dimensional vision on laparoscopic performance. *J Endourol* 28:261–266
 20. Kowalewski TM, White LW, Lendvay TS, Jiang IS, Sweet R, Wright A, Hannaford B, Sinanan MN (2014) Beyond task time: automated measurement augments fundamentals of laparoscopic skills methodology. *J Surg Res* 192:329–338
 21. Steinemann DC, Limani P, Clavien PA, Breitenstein S (2013) Internal retraction in single-port laparoscopic cholecystectomy: initial experience and learning curve. *Minim Invasive Ther Allied Technol* 22:171–176
 22. Hennessey IA, Hewett P (2013) Construct, concurrent, and content validity of the eoSim laparoscopic simulator. *J Laparoendosc Adv Surg Tech A* 23:855–860
 23. Boza C, Varas J, Buckel E, Achurra P, Devaud N, Lewis T, Aggarwal R (2013) A cadaveric porcine model for assessment in laparoscopic bariatric surgery—a validation study. *Obes Surg* 23:589–593
 24. Palter VN, Graafland M, Schijven MP, Grantcharov TP (2012) Designing a proficiency-based, content validated virtual reality curriculum for laparoscopic colorectal surgery: a Delphi approach. *Surgery* 151:391–397
 25. Pagador JB, Sanchez-Margallo FM, Sanchez-Peralta LF, Sanchez-Margallo JA, Moyano-Cuevas JL, Enciso-Sanz S, Uson-Gargallo J, Moreno J (2012) Decomposition and analysis of laparoscopic suturing task using tool-motion analysis (TMA): improving the objective assessment. *Int J Comput Assist Radiol Surg* 7:305–313
 26. Yamaguchi S, Yoshida D, Kenmotsu H, Yasunaga T, Konishi K, Ieiri S, Nakashima H, Tanoue K, Hashizume M (2011) Objective assessment of laparoscopic suturing skills using a motion-tracking system. *Surg Endosc* 25:771–775
 27. Strickland A, Fairhurst K, Lauder C, Hewett P, Maddern G (2011) Development of an ex vivo simulated training model for laparoscopic liver resection. *Surg Endosc* 25:1677–1682
 28. Hiemstra E, Chmarra MK, Dankelman J, Jansen FW (2011) Intracorporeal suturing—economy of instrument movements using a box trainer model. *J Minim Invasive Gynecol* 18:494–499
 29. Zheng B, Cassera MA, Martinec DV, Spaun GO, Swannstrom LL (2010) Measuring mental workload during the performance of advanced laparoscopic tasks. *Surg Endosc* 24:45–50
 30. Lin DW, Romanelli JR, Kuhn JN, Thompson RE, Bush RW, Seymour NE (2009) Computer-based laparoscopic and robotic surgical simulators: performance characteristics and perceptions of new users. *Surg Endosc* 23:209–214
 31. Kroeze SG, Mayer EK, Chopra S, Aggarwal R, Darzi A, Patel A (2009) Assessment of laparoscopic suturing skills of urology residents: a pan-European study. *Eur Urol* 56:865–872
 32. Botden SM, de Hingh IH, Jakimowicz JJ (2009) Meaningful assessment method for laparoscopic suturing training in augmented reality. *Surg Endosc* 23:2221–2228
 33. Zheng B, Denk PM, Martinec DV, Gatta P, Whiteford MH, Swannstrom LL (2008) Building an efficient surgical team using a bench model simulation: construct validity of the Legacy Inanimate System for Endoscopic Team Training (LISETT). *Surg Endosc* 22:930–937
 34. Dayan AB, Ziv A, Berkenstadt H, Munz Y (2008) A simple, low-cost platform for basic laparoscopic skills training. *Surg Innov* 15:136–142
 35. Stefanidis D, Scerbo MW, Korndorffer JR Jr, Scott DJ (2007) Redefining simulator proficiency using automaticity theory. *Am J Surg* 193:502–506
 36. Dubrowski A, Larmer JC, Leming JK, Brydges R, Carnahan H, Park J (2006) Quantification of process measures in laparoscopic suturing. *Surg Endosc* 20:1862–1866
 37. Broe D, Ridgway PF, Johnson S, Tierney S, Conlon KC (2006) Construct validation of a novel hybrid surgical simulator. *Surg Endosc* 20:900–904
 38. Van Sickle KR, McClusky DA 3rd, Gallagher AG, Smith CD (2005) Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. *Surg Endosc* 19:1227–1231
 39. Duffy AJ, Hogle NJ, McCarthy H, Lew JI, Egan A, Christos P, Fowler DL (2005) Construct validity for the LAPSIM laparoscopic surgical simulator. *Surg Endosc* 19:401–405
 40. Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A (2004) Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. *Surg Endosc* 18:1608–1612
 41. Figert PL, Park AE, Witzke DB, Schwartz RW (2001) Transfer of training in acquiring laparoscopic skills. *J Am Coll Surg* 193:533–537
 42. Poursartip B, LeBel ME, Patel R, Naish M, Trejos AL (2017) Analysis of energy-based metrics for laparoscopic skills assessment. *IEEE Trans Biomed Eng* 65:1532–1542
 43. Botden SM, Buzink SN, Schijven MP, Jakimowicz JJ (2007) Augmented versus virtual reality laparoscopic simulation: what is the difference? A comparison of the ProMIS augmented reality laparoscopic simulator versus LapSim virtual reality laparoscopic simulator. *World J Surg* 31:764–772
 44. Kobayashi SA, Jamshidi R, O’Sullivan P, Palmer B, Hirose S, Stewart L, Kim EH (2011) Bringing the skills laboratory home: an affordable webcam-based personal trainer for developing laparoscopic skills. *J Surg Educ* 68:105–109
 45. Trudeau MO, Nasr A, Carrillo B, Gerstle JT, Azzie G (2015) Construct validity and educational role for motion analysis in a laparoscopic trainer. *Surg Endosc* 29:2491–2495
 46. Sharma M, Macafee D, Pranesh N, Horgan AF (2012) Construct validity of fresh frozen human cadaver as a training model in minimal access surgery. *JLS* 16:345–352
 47. Kowalewski K-F, Hendrie JD, Schmidt MW, Garrow CR, Bruckner T, Proctor T, Paul S, Adiguzel D, Bodenstedt S, Erben A, Kennigott H, Erben Y, Speidel S, Muller-Stich BP, Nickel F (2017) Development of a sensor-and expert model based training device for suturing and knot tying in laparoscopic surgery. *Surg Endosc* 31:2155–2165
 48. Horeman T, Rodrigues SP, Jansen FW, Dankelman J, Van Den Dobbelen JJ (2010) Force measurement platform for training

- and assessment of laparoscopic skills. *Surg Endosc Other Interv Tech* 24:3102–3108
49. Rosser JC Jr, Liu X, Jacobs C, Choi KM, Jalink MB, Ten Cate Hoedemaker HO (2017) Impact of Super Monkey Ball and Underground video games on basic and advanced laparoscopic skill training. *Surg Endosc* 31:1544–1549
 50. Zdichavsky M, Krautwald M, Feilitzsch MV, Wichmann D, Konigsrainer A, Schurr MO (2016) Laparoscopic gastro-jejunal anastomosis using novel r2 deflectable instruments in an ex vivo model. *Minim Invasive Ther Allied Technol* 25:91–98
 51. Uemura MYM, Tomikawa M, Obata S, Souzaki R, Ieiri S, Ohuchida K, Matsuoka N, Katayama T, Hashizume M (2015) Objective assessment of the suture ligation method for the laparoscopic intestinal anastomosis model using a new computerized system. *Surg Endosc* 29(2):444–452
 52. Veneziano D, Poniatowski LH, Reihnsen TE, Sweet RM (2016) Preliminary evaluation of the SimPORTAL major vessel injury (MVI) repair model. *Surg Endosc* 30:1405–1412
 53. Keyser EJ, Derossis AM, Antoniuk M, Sigman HH, Fried GM (2000) A simplified simulator for the training and evaluation of laparoscopic skills. *Surg Endosc* 14:149–153
 54. Xeroulis G, Dubrowski A, Leslie K (2009) Simulation in laparoscopic surgery: a concurrent validity study for FLS. *Surg Endosc* 23:161–165
 55. Bahsoun AN, Malik MM, Ahmed K, El-Hage O, Jaye P, Dasgupta P (2013) Tablet based simulation provides a new solution to accessing laparoscopic skills training. *J Surg Educ* 70:161–163
 56. Oostema JA, Abdel MP, Gould JC (2008) Time-efficient laparoscopic skills assessment using an augmented-reality simulator. *Surg Endosc* 22:2621–2624
 57. Sleiman Z, Atallah E, Rassi E, Sarkis R, Khazzaka A (2017) Validation study of a portable home trainer using a pad for laparoscopic practice. *Surg Innov* 24:284–288
 58. Yeung C, Carrillo B, Pope V, Hosseinpour S, Gerstle JT, Azzie G (2017) Video assessment of laparoscopic skills by novices and experts: implications for surgical education. *Surg Endosc* 31:3883–3889
 59. Anton NE, Sawyer JM, Korndorffer JR Jr, DuCoin CG, McRary G, Timsina LR, Stefanidis D (2018) Developing a robust suturing assessment: validity evidence for the intracorporeal suturing assessment tool. *Surgery* 163:560–564
 60. Bilgic E, Takao M, Kaneva P, Satoshi E, Takao T, Watanabe Y, McKendy KM, Feldman LS, Vassiliou MC (2018) Development of a model for the acquisition and assessment of advanced laparoscopic suturing skills using an automated device. *Surg Innov* 25:286–290
 61. Bilgic E, Watanabe Y, Nepomnayshy D, Gardner A, Fitzgibbons S, Ghaderi I, Alseidi A, Stefanidis D, Paige J, Seymour N, McKendy KM, Birkett R, Whitley J, Kane E, Anton NE, Vassiliou MC, Simulation Committee of the Association for Surgical E (2017) Multicenter proficiency benchmarks for advanced laparoscopic suturing tasks. *Am J Surg* 213:217–221
 62. McKendy KM, Watanabe Y, Bilgic E, Enani G, Munshi A, Lee L, Feldman LS, Fried GM, Vassiliou MC (2017) Establishing meaningful benchmarks: the development of a formative feedback tool for advanced laparoscopic suturing. *Surg Endosc* 31:5057–5065
 63. Nepomnayshy D, Whitley J, Birkett R, Delmonico T, Ruthazer R, Sillin L, Seymour NE (2015) Evaluation of advanced laparoscopic skills tasks for validity evidence. *Surg Endosc* 29:349–354
 64. Poursartip B, Lebel ME, Patel RV, Naish MD, Trejos AL (2018) Analysis of energy-based metrics for laparoscopic skills assessment. *IEEE Trans Biomed Eng* 65:1532–1542
 65. Sawyer JM, Anton NE, Korndorffer JR Jr, DuCoin CG, Stefanidis D (2018) Time crunch: increasing the efficiency of assessment of technical surgical skill via brief video clips. *Surgery* 163:933–937
 66. Takayasu K, Yoshida K, Mishima T, Watanabe M, Matsuda T, Kinoshita H (2018) Upper body position analysis of different experience level surgeons during laparoscopic suturing maneuvers using optical motion capture. *Am J Surg* 21:12–16
 67. Watanabe Y, McKendy KM, Bilgic E, Enani G, Madani A, Munshi A, Feldman LS, Fried GM, Vassiliou MC (2016) New models for advanced laparoscopic suturing: taking it to the next level. *Surg Endosc* 30:581–587
 68. Wee J, Azzie G, Drake J, Gerstle JT (2018) Proof of concept study: investigating force metrics of an intracorporeal suturing knot task. *J Laparoendosc Adv Surg Tech A* 28:899–905
 69. Yamaguchi T, Nakamura R (2018) Laparoscopic training using a quantitative assessment and instructional system. *Int J Comput Assist Radiol Surg* 13:1453–1460
 70. Vassiliou MC, Ghitulescu GA, Feldman LS, Stanbridge D, Lefondré K, Sigman HH, Fried GM (2006) The MISTELS program to measure technical skill in laparoscopic surgery: evidence for reliability. *Surg Endosc* 20:744–747
 71. Ritter ME, Scott DJ (2007) Design of a proficiency-based skills training curriculum for the fundamentals of laparoscopic Surgery. *Surg Innov* 14:107–112
 72. Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, Andrew CG (2004) Proving the value of simulation in laparoscopic surgery. *Ann Surg* 240:518–528
 73. Fried GM (2008) FLS assessment of competency using simulated laparoscopic tasks. *J Gastrointest Surg* 12:210–212
 74. Mason JD, Ansell J, Warren N, Torkington J (2013) Is motion analysis a valid tool for assessing laparoscopic skill? *Surg Endosc* 27:1468–1477
 75. Chmarra MK, Grimbergen CA, Dankelman J (2007) Systems for tracking minimally invasive surgical instruments. *Minim Invasive Ther Allied Technol* 16:328–340
 76. Pugh CM, Arafat FO, Kwan C, Cohen ER, Kurashima Y, Vassiliou MC, Fried GM (2015) Development and evaluation of a simulation-based continuing medical education course: beyond lectures and credit hours. *Am J Surg* 210:603–609

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