Development and validation of the objective assessment of robotic suturing and knot tying skills for chicken anastomotic model

Stefano Puliatti, Elio Mazzone, Marco Amato, Ruben De Groote, Alexandre Mottrie & Anthony G. Gallagher

ONLINE FIRST

Surgical Endoscopy
And Other Interventional Techniques
Official Journal of the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) and European Association for Endoscopic Surgery (EAES)

ISSN 0930-2794

Surg Endosc
DOI 10.1007/s00464-020-07918-5

Springer

Surgical Endoscopy
https://doi.org/10.1007/s00464-020-07918-5

Development and validation of the objective assessment of robotic suturing and knot tying skills for chicken anastomotic model

Stefano Puliatti[1,2,3] · Elio Mazzone[4,5] · Marco Amato[1,2,3] · Ruben De Groote[1,2] · Alexandre Mottrie[1,2] · Anthony G. Gallagher[1,6,7]

Received: 10 June 2020 / Accepted: 17 August 2020

Abstract

Background To improve patient safety, there is an imperative to develop objective performance metrics for basic surgical skills training in robotic surgery.

Objective To develop and validate (face, content, and construct) the performance metrics for robotic suturing and knot tying, using a chicken anastomotic model.

Design, setting and participants Study 1: In a procedure characterization, we developed the performance metrics (i.e., procedure steps, errors, and critical errors) for robotic suturing and knot tying, using a chicken anastomotic model. In a modified Delphi panel of 13 experts from four EU countries, we achieved 100% consensus on the five steps, 18 errors and four critical errors (CE) of the task.

Study 2: Ten experienced surgeons and nine novice urology surgeons performed the robotic suturing and knot tying chicken anastomotic task. The mean inter-rater reliability for the assessments by two experienced robotic surgeons was 0.92 (95% CI, 0.9–0.95). Novices took 18.5 min to complete the task and experts took 8.2 min. ($p = 0.00001$) and made 74% more objectively assessed performance errors than the experts ($p = 0.000343$).

Conclusions We demonstrated face, content, and construct validity for a standard and replicable basic anastomotic robotic suturing and knot tying task on a chicken model.

Patient summary Validated, objective, and transparent performance metrics of a robotic surgical suturing and knot tying tasks are imperative for effective and quality assured surgical training.

Keywords Surgical training · Proficiency-based metrics · Face · content and construct validation

There were more than 4500 Da Vinci robotic systems (Intuitive Surgical Inc., Sunnyvale, CA, USA) operating around the world in 2018 and are currently the most widely used robotic system [1]. There are, however, a number of new robotic surgical systems being introduced into the market place over the next 2 years. The robotic technique has shown advantages over the laparoscopic approach especially a shorter learning curve [2]. However, to date, the implementation of basic surgical training pathways for robotic surgery has been suboptimal and, as a consequence, many novice robotic surgeons start their practice in the operating room (OR) with limited or no experience. The need to introduce surgeons in the OR when they reach the proficiency level in technical skills is one of the most critical problems that the surgical community is trying to solve [3, 4]. The inadequate preparation of robotic surgeons can result in higher risk of adverse events during procedures which in turn may have

First authorship shared by Stefano Puliatti and Elio Mazzone.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00464-020-07918-5) contains supplementary material, which is available to authorized users.

✉ Marco Amato
marcohz92@gmail.com

1 ORSI Academy, Melle, Belgium

2 Department of Urology, OLV, Aalst, Belgium

3 Department of Urology, University of Modena and Reggio Emilia, Modena, Italy

4 Division of Oncology/Unit of Urology, URI, IRCCS Ospedale San Raffaele, Milan, Italy

5 Vita-Salute San Raffaele University, Milan, Italy

6 Faculty of Life and Health Sciences, Ulster University, Derry, Northern Ireland, UK

7 Faculty of Medicine, KU Leuven, Leuven, Belgium

Published online: 28 August 2020

negative implications for patients' outcomes [2, 5]. There is therefore a need for the development of surgical training programs which are based on objective, transparent, fair and validated performance metrics. This approach ensures that novices are well trained in the skills laboratory, before they operate on real patients. Proficiency-based progression (PBP) is a specific training methodology that has shown in several prospective, randomized, and double-blind studies to produce better training outcomes compared to traditional training methods [6–12]. This methodology has also been shown to be effective in improving clinical outcomes [8].

The first step in creating a proficient robotic surgeon is the acquisition of basic surgical skills, such as suturing, knotting, coagulating, and dissecting [13]. Some training models mimic the real anatomical environment in which a procedure will be performed and provide trainees with the opportunity to learn and practice their skills before starting their practice on real cases in the operating room. Numerous dry-lab, wet-lab, and virtual reality simulation models have been described which purport to achieve this effect [14]. The "Venezuelan chicken model" seems to be a good model for trainees to learn robotic suturing, anastomosis, and knot tying (for all surgical disciplines) [14–16]. In order to establish a structured PBP training pathway, objective performance metrics for optimal and suboptimal performance need to be developed. In this study, we report the development and validation of performance metrics for robotic suturing and knot tying using the Venezuelan chicken model.

Materials and methods

Subjects

The performance metrics were presented to a panel of 13 experts from four different countries (Table 1) in a modified online Delphi process. The criteria used to select the Delphi consensus participants were having (1) experience in robotic surgery, (2) peer-reviewed publications on training and robotic surgery, and (3) participated in training courses in which the "Venezuelan chicken model" had been used. An initial list of 20 possible eligible participants has been drawn up. The 20 candidates were invited by email, 13 of them agreed to participate in the Delphi Consensus.

Ten experienced surgeons [from Belgium ($n=5$), Netherlands ($n=3$), India ($n=1$), Spain ($n=1$)] performed > 300 robotic procedures and 9 novice surgeons [from Italy ($n=3$), Poland ($n=1$), Portugal ($n=1$), Norway ($n=1$), Germany ($n=1$), India ($n=1$), England ($n=1$)]. The novice surgeons who participated in the construct validity study had completed/performed < 5 full robotic procedures and the experienced group reported that they had completed between 300 and 2000 robotic procedures. All the novices included in the study participated in the Robotic Skills Course (CC-ERUS) for Robot-Assisted Radical Prostatectomy [17], or in an equivalent 1-week program with the same tasks and duration. Median age of experienced and novice surgeons was, respectively, 49 and 37 years old, respectively.

Table 1 Demographic characteristics of 13 international participants of the Delphi consensus meeting. (*RALP* robot-assisted laparoscopic prostatectomy, *RAKT* robot-assisted kidney transplantation, *RARC* robot-assisted radical cystectomy)

Experts	Country	Years of experience	Expertise
1	Italy	9	RALP
2	Italy	4	RALP
3	Italy	6	RALP
4	Italy	10	RALP
5	Italy	12	RALP
6	Belgium	12	RAKT, RARC
7	Belgium	15	RALP
8	Belgium	10	RALP
9	Poland	13	RALP
10	Italy	5	RALP
11	Belgium	20	RALP
12	Portugal	15	RALP, RAPN
13	Belgium	10	RARC

Model preparation, robot set-up, and anastomosis technique

The preparation of the chicken model was standardized. The chicken is provided after removing all abdominal organs, except for 6–8 cm of the cloaca (starting measurements from the anus) and the stomach (Supplementary Figures a, b). The chicken can then be frozen and preserved until the day of its use for training. The thawing process involves running water at 60° on the surface and inside the chicken for about 1 min and then leaving the chicken at room temperature for about 1 h. The preparation of the chicken starts by opening chicken legs, followed by cutting and removal of the lower and upper anterior abdominal wall. Finally, a 12 Ch Foley latex catheter is placed in the cloaca (Supplementary Figures c–e). The cost of the model varies from 4 to 7 euros. Ideally, the cloaca must overlap with the stomach for 1 cm. Shorter or longer overlap may impede the correct execution of the urethro-vesical anastomosis training.

The suturing and knot tying chicken model can be completed with any surgical robot model currently available on the market and we also propose a standardized process for the robot positioning. An appropriate training location with sufficient spaces allowing for easy robot patient cart movement is suggested. The table where the chicken model is placed should be fixed. The patient cart is positioned in order to maintain the camera trocar at a distance of 18–20 cm from the target (Supplementary Figures f, g). Robotic trocars must be positioned at a distance of 8–10 cm apart (Supplementary Figures h). Two large needle driver and one Prograsp™ forceps are used. A pelvic box trainer is not necessary for robot and trocar docking.

Two Polysorb™ sutures of 10 cm each, knotted at the end with four knots, are used to perform the task (Supplementary Figures i). The anastomosis is executed according to the technique described by Roland F. Van Velthoven et al. [18] (Supplementary Figures l).

Face and content validity (Delphi consensus)

This study received expedited Institutional Review Board approval from Onze Lieve Vrouw Hospital, Aalst, Belgium (OLV – studienummer: 2019/093). Chicken anastomosis characterization was performed in three face-to-face meetings. Four surgeons (AM, RDG, EM, SP) and a Behavioral Scientist (AGG) formed the procedure characterization group. Procedure characterization methods are described elsewhere [6, 19–22].

The online Delphi meeting was coordinated with the "Microsoft Teams" web platform on April 27, 2020. At the start of the meeting, the concepts of 'PBP' were outlined. The procedure metrics for a reference approach to the suturing and knot tying in a chicken anastomotic model were presented. Steps, errors, and critical errors (CE) were outlined and discussed by the Delphi panel. Following this discussion, the proposed metrics were edited in real time and voted on to establish the level of consensus on the metrics.

Construct validity

For the construct validation, we compared the objectively scored performance of ten very experienced and nine novice robotic surgeons, using the final version of the anastomotic model with the agreed performance metrics. Two experienced robotic surgeons were trained to score the metrics until they consistently achieved higher than 0.8 inter-rater reliability (IRR). Reviewer training (detailed methodology described elsewhere) was initiated with an 8-h meeting, during which time each metric was studied in detail [23]. Multiple video examples of chicken anastomosis tasks were shown to illustrate each particular metric. The discussion helped to clarify how each step and error was to be scored, including the nuances and conventions to be used. Full-length practice videos were then independently scored by each of the reviewers, and the scores were tabulated. The differences and discrepancies between the reviewers were compared and discussed, seeking consensus in scoring. Practice video scoring continued until the reviewers achieved an' *IRR* (agreements/agreements + disagreements) ≥ 0.8, consistently. Only then did reviewers progress to scoring study videos [23].

The video reviewers remained blinded as to the identity of the operator and their status (i.e., experienced or novice surgeon). The IRR was calculated for all participants (i.e., agreements/agreements + disagreements) [24].

Statistical methods

In a pilot for Study 2 using the metrics, we assessed the performance of three novices and three experts. We established that the metrics were scorable reliably (mean inter-rater reliability = 0.94). The observed mean error rate was 15.7 [standard deviation (SD) = 2.1] for the novice group and for the experts = 11.7 (SD = 3.5). For statistical power, assuming an Alpha Error Level of Level 0.05, and Beta Error Level or Statistical Power [1 − Beta] of 20% a sample size = 6 was required in each group for the observed difference to reach statistical significance. In a similar analysis for the Time metric with the same assumptions as the Error metric we found that [mean time; novices = 16.1 (SD = 4.2) and experts = 9.4 (SD = 2.11)] a sample size $n = 3$ was required in each group for the observed difference to reach statistical significance.

Changes in the number of metric units before and after the Delphi meeting were compared for statistical significance with Wilcoxon signed rank tests. The relationship between the number of metric units before and after was assessed with Pearson's product moment correlation coefficient. Differences in performance between the two groups (experienced vs. novice operators) were assessed using one-factor analysis of variance (ANOVA) and Chi-square tests. Statistical analysis was performed with SPSS (IBM SPSS statistical software program, version 24; IBM, Armonk, NY, USA) and with the R software (version 3.5.1; https://www.r-project.org/). Statistical significance for all analyses was defined as $p < 0.05$.

Results

Study 1: Delphi consensus meeting

The number of steps, errors, and CE before and after the Delphi panel meeting are shown in Table 2. Overall, 100% (13/13) of the panel was fellowship trained; median age of panel experts was 37 years old. Only one notable change and edit to the metrics was proposed by the Delphi panel. Table 2 summarizes the results obtained during the Delphi consensus. One CE was added to the initially proposed metrics. After this modification, all the metrics were accepted with 100% consensus by the participants. Metric units before

Table 2 Steps, errors, and critical errors, before and after the modified Delphi consensus

	N. before Delphi	*N*. after Delphi	Agreement (%)
1. Steps	5	5	100
2. Suturing operative errors	12	12	100
3. Knotting operative errors	5	5	100
4. "Fail to progress" (FTP) definition	1	1	100
5. "Phase completed according to our definition" (CMP)	1	1	100
6. Critical errors	3 (1)*	4	100

()*modifications in parentheses

and after the Delphi were strongly positively correlated ($r = 0.995$, $p < 0.000$).

The final template of the metrics table developed for the suturing and knot tying chicken model is reported in Fig. 1. The procedure was divided into five steps: posterior, left, right, anterior walls, and knotting. The steps are defined precisely, based on the number of bites performed on the cloaca (which corresponds to the human urethra) and stomach (which corresponds to the human bladder). Scoring commences when the subject engages the first bite on the chicken bladder. The precise definition of each step allows the avoidance of misunderstandings in assigning errors to one phase rather than another. Overall, 12 suturing operative errors and five knotting operative errors were identified and carefully defined. Some suturing operative errors are not applicable to the knotting step, and vice versa, the knotting operative errors are specific and, therefore, applicable only to the knotting step. Suturing and knotting operative errors can be scored only once in each single step. The failure to progress (FTP) was also defined. It is scored when no progress is made in a specific step of the anastomosis for one complete minute and it may be scored more than once, i.e., each time the subject fails to make progress in a complete one-minute period. Leakage of the anastomosis, breakage of the needle or suture, and catheter fixation during anastomosis are considered critical errors and the task must be completed in 40 min.

Study 2: construct validity assessment

Overall, ten experienced and nine novice (but trained) surgeons were evaluated. The mean and 95% confidence intervals (CI) amount of time to complete the task by the two groups are shown in Fig. 2. On average, novices s took 18.5 min to complete the task and experts took 8.2 min. The performance variability to complete the task was > 2 times greater for the novices in comparison to the experts. When compared with ANOVA, this difference was statistically significant [F (1, 16) = 40.02, $p = 0.00001$].

A similar pattern was observed for the mean number of errors made during the performance of the chicken anastomosis model. Novices made 74% more objectively assessed performance errors than the experts. This difference was also statistically significant [F (1, 17) = 19.9, $p = 0.000343$] (Fig. 3).

Novices were more than 12½ times more likely to demonstrate failure to progress (Fig. 4) than the expert group. The novices also showed substantial higher performance variability than the experts, but the mean difference did not reach statistically significance [F (1, 17) = 2.42, $p = 0.138$].

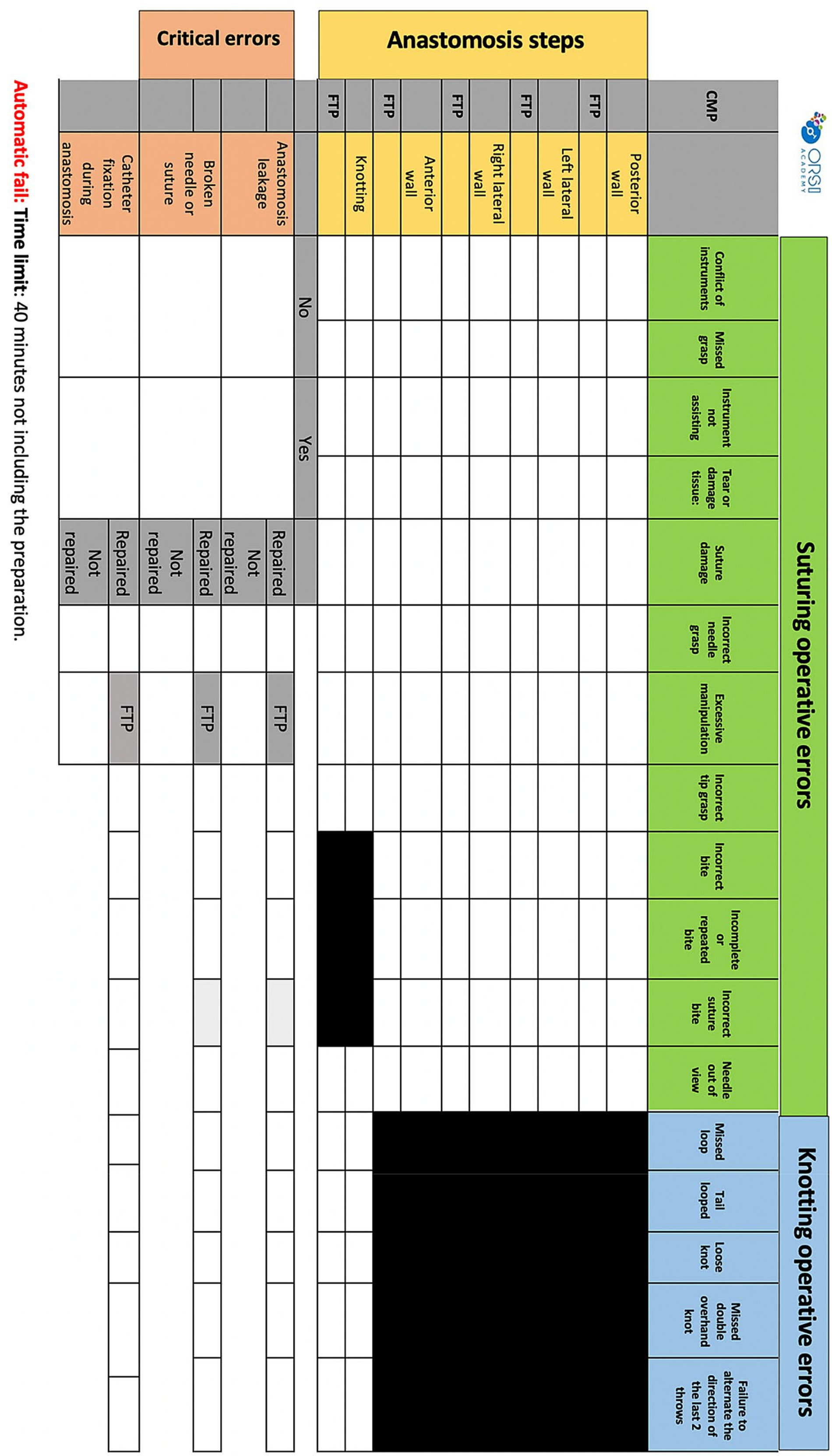
When assessed, six (67%) of the nine novices had anastomosis leakage at the end of the task. Only two (20%) of the expert group had anastomotic leakage. When compared for significance with Chi-square, this pattern was found to be statistically significant (Chi-square = 4.232, df = 1, $p = 0.04$). Overall, mean inter-rater reliability for all of the chicken anastomosis models was 0.92 (95% CI, 0.9–0.95). The mean IRR for the experts was slightly higher (i.e., 0.93) than for novices (i.e., 0.91). No assessment fell below the 0.8 IRR level.

Discussion

In the last 20 years, robotic surgery has developed enormously, showing exponential growth for minimally invasive surgical procedures worldwide [3, 25]. This cutting-edge technology has demonstrated efficacy in improved patient outcomes and safety [2]. However, it is of primary importance to combine technological development with an adequate preparation of robotic surgeons to achieve high standards of care [13]. In this context, training curricula are essential to deliver optimal pathways for robotic surgery trainees [17, 26, 27]. In order to characterize the entire training process from the skills lab to the OR, different skill levels may be defined and learned by the trainees. Specifically, dry-lab technical skills training represents a fundamental step to improve manual dexterity and acquisition of basic robotic surgical skills.

The chicken model has been successfully proposed as a dry-lab model for basic skills training in multiple surgical specialties [14–16]. The chicken model has several

Fig. 1 The scoresheet for the chicken anastomotic model



CMP		Suturing operative errors: Conflict of instruments	Missed grasp	Instrument not assisting	Tear or damage tissue:	Suture damage	Incorrect needle grasp	Excessive manipulation	Incorrect tip grasp	Incorrect bite	Incomplete or repeated bite	Incorrect suture bite	Needle out of view	Knotting operative errors: Missed loop	Tail looped	Loose knot	Missed double overhand knot	Failure to alternate the direction of the last 2 throws
Anastomosis steps	Posterior wall																	
FTP																		
	Left lateral wall																	
FTP																		
	Right lateral wall																	
FTP																		
	Anterior wall																	
FTP																		
	Knotting																	
FTP																		
		No		Yes														
Critical errors	Anastomosis leakage					Repaired		FTP										
						Not repaired												
	Broken needle or suture					Repaired		FTP										
						Not repaired												
	Catheter fixation during anastomosis					Repaired		FTP										
						Not repaired												

Automatic fail: Time limit: 40 minutes not including the preparation.

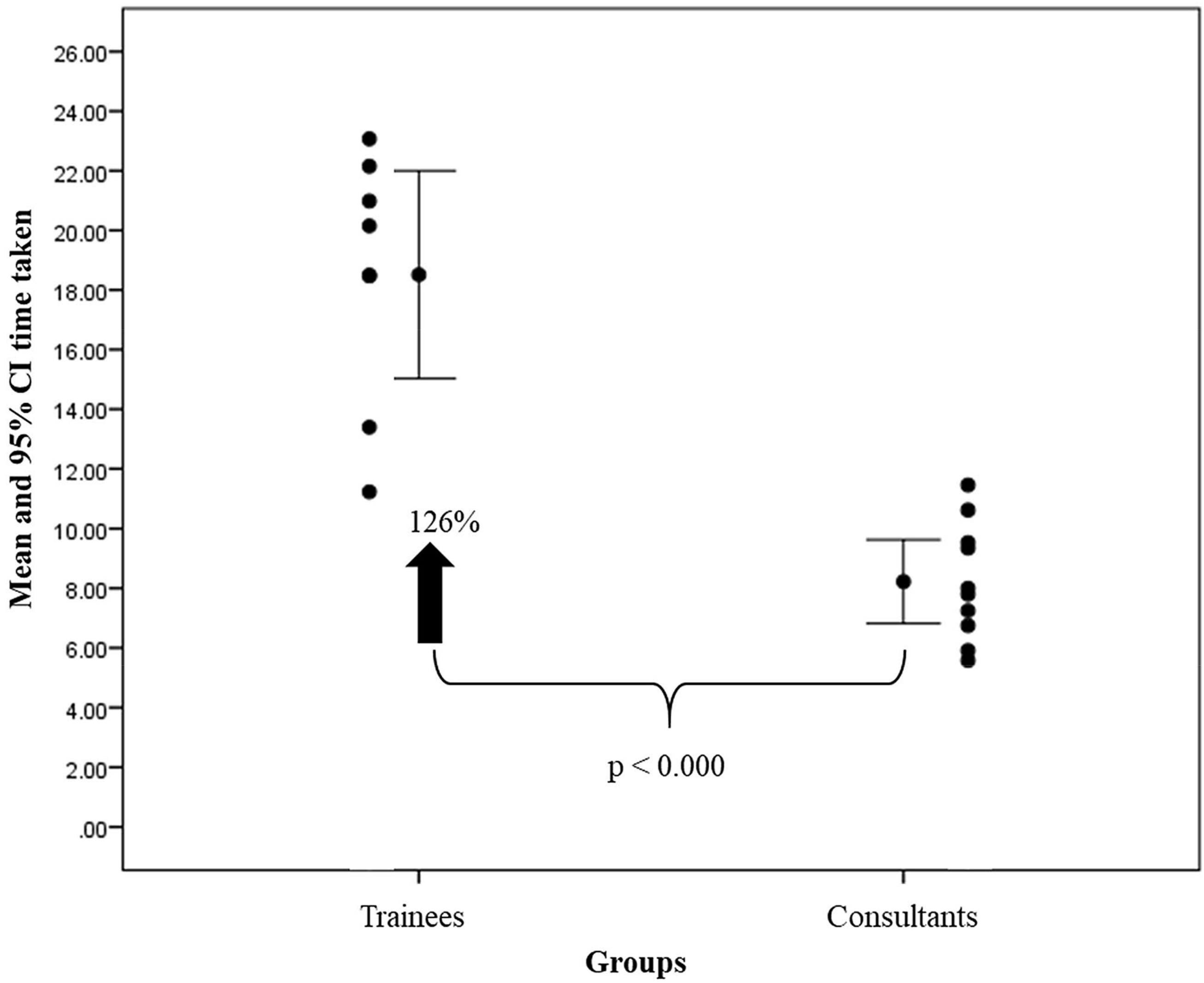


Fig. 2 Individual scores, the mean, and 95% confidence intervals of the amount of time it took the two groups to complete the chicken anastomosis task

advantages: it is inexpensive, widely available, and easy to prepare, it is a good model for the urethro-vesical anastomosis training, it is possible to introduce a catheter and performing the leakage test, and it takes the trainee approx. 20 min to complete the task. There is, however, a lack of standardized and objective performance metrics based on this model which would allow implementation of PBP training on suturing and knot tying tasks.

Performance metrics are an integral part of PBP training. The metrics explicitly characterize optimal (i.e., procedure steps) and suboptimal (i.e., procedure errors and critical errors) task performance [19–21]. They are used to give the trainee formative feedback on their performance and have been demonstrated to significantly enhanced the acquisition of skills [6, 9]. Performance metrics are also used to quantitatively define proficiency benchmarks which trainees must attain before training is deemed completed. It has been shown in prospective, randomized, and blinded clinical studies that trainees who receive the exact same curriculum, for the same period of time, using the same training models and taught by the same level faculty but without the metric-based feedback and proficiency benchmarks perform significantly worse than PBP trainees who did have the metric-based feedback and the proficiency benchmark requirements [6–9, 11, 12, 28–32]. Based on this premise, we developed and validated objective performance metrics on the ORSI suturing and knot tying chicken model. To achieve this goal, a two-phase study was conducted.

First, we achieved consensus among very experienced surgeons on the key steps and errors which characterize this model. Using a modified Delphi process complete consensus was reached on the proposed steps and errors

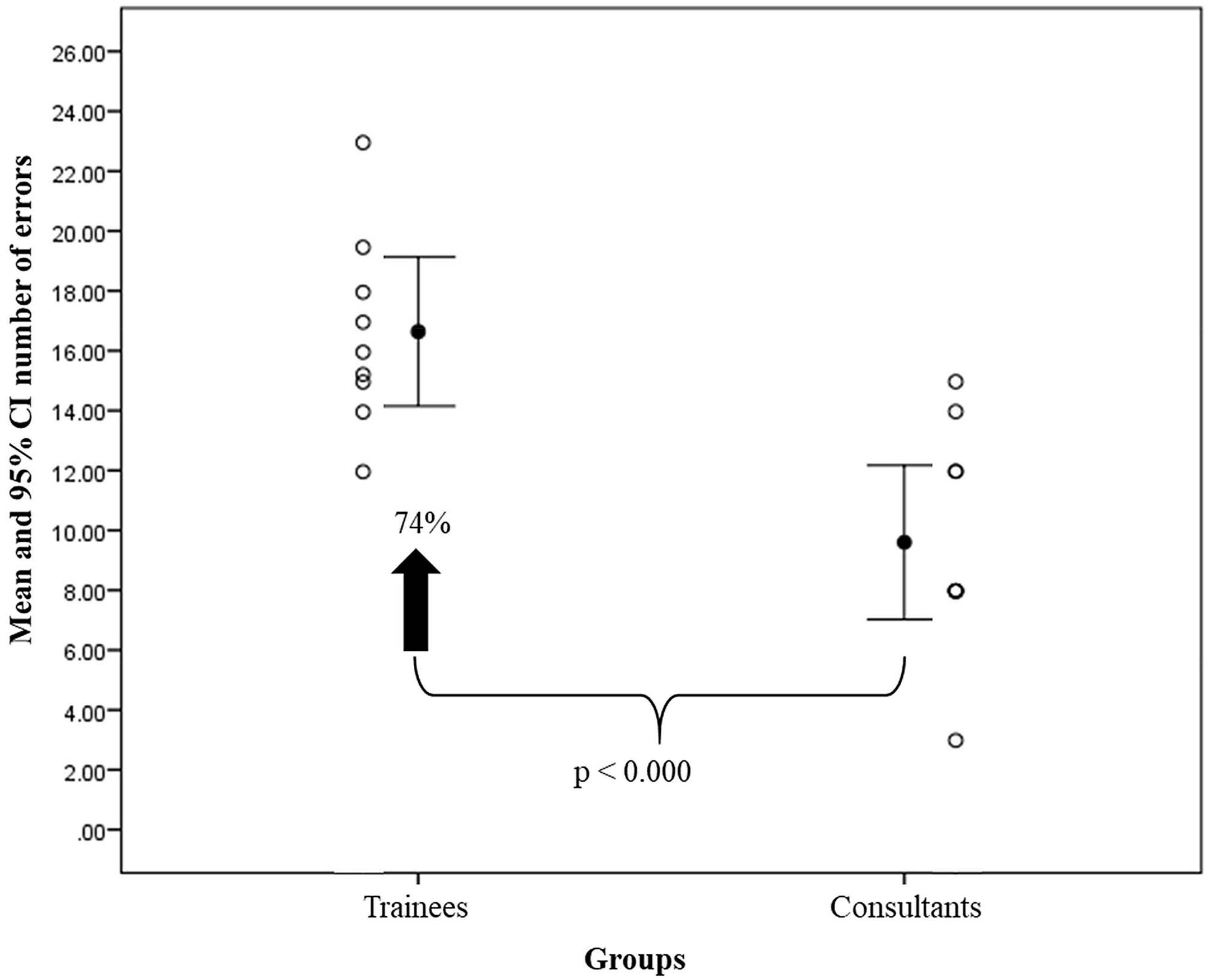


Fig. 3 Individual scores, the mean, and 95% confidence intervals of the number of performance errors made by the two groups when completing the chicken anastomosis task

and critical errors. At the Delphi meeting, one critical error (i.e., suturing of the catheter into the anastomosis) was added to the metrics template. The 100% agreement obtained may be explained by preparty efforts on the detailed deconstruction of a full anastomosis task which resulted in the proposed template. This template closely emulates previously validated anastomotic models proposed for different specialties. Van Sickle et al. [11] developed a comparable template for a Laparoscopic Nissen fundoplication. His approach represented a good format to set up the steps and errors of a new basic robotic suturing and knot tying task model. Moreover, through discussion between experts, we advanced and improved all the definitions in order to optimize their application on the newly developed suturing and knot tying model.

Time and failure to progress (FTP) in task performance accurately discriminated between expert and novice groups, demonstrating construct validity for 'process' metrics. These findings are in line with those previously reported by Van Sickle et al. [11]. In the study reported here, time and FTP reliably differentiated between practitioner experience levels, where a difference of 40% in performance time was observed. It should, however, be noted that process metrics such as time may not reliably assess performance quality.

Metrics of performance quality (errors and critical errors) also demonstrated construct validity. Performance quality metrics are a good basis to define an objective quality assessment of a surgical task. Novices made more than double the number of errors than the expert group. These results replicate those reported by Van Sickle et al.

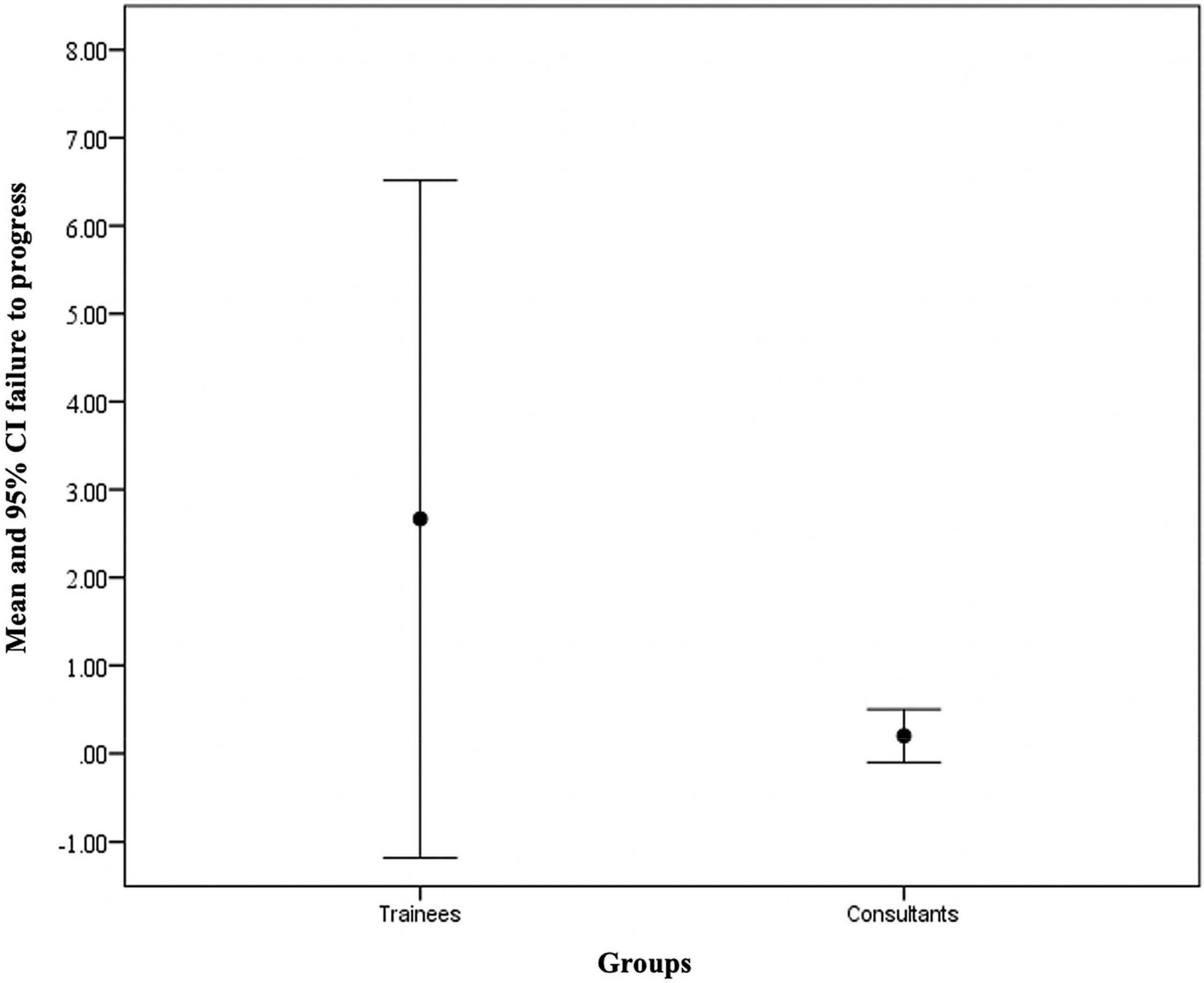


Fig. 4 The mean and 95% confidence intervals of the number of 'failure to progress' units scored for the two groups when completing the chicken anastomosis task

[11], where a reduction in errors of 30% was reported using a similar training model. Similarly, Pedowitz et al. [10] reported a reduction of 60% in basic arthroscopic dry-lab knot tying performance.

Taken together, we demonstrated face, content, and construct validity of a novel dry-lab basic suturing skills model by using a modified Delphi consensus process and subsequent prospective construct validation of the identified metrics. Defining and validating objective quantitative performance metrics are crucial steps to enable a PBP training pathway [19]. This model thus establishes the basis to implement a PBP methodology for basic skills training.

Limitations of the current analyses are the relatively small sample size and the variability of novices' performances. However, the prospective nature of the study and the adequate sample size definition allowed us to demonstrate the construct validity of the newly developed metrics.

The metric-based characterization for the task used in this study was developed on a specific robotic surgical system (DaVinci™). Thus, the anastomotic task and its related metrics may not be replicable in different robotic platforms, where dedicated development and validation of performance metrics are necessary. Nonetheless, given the similarity of these metrics with those previously reported in laparoscopic surgery, it is plausible that most of the proposed metrics may be partially or fully applicable to other robotic systems.

The effort required to assess performance with this level of robustness and rigor is considerable. That said, we are very aware of the possibility that our task, assessment methods and metrics could in the future be used for high stakes assessment of novices. Therefore, only the highest standard of scientific methods is acceptable. We are currently working with a number of researchers to investigate the possibility of the metrics being scored by a machine learning/artificial intelligence system. The building blocks for this approach are the validated performance metrics. This approach is, however, in its infancy.

Conclusions

After a procedure characterization, performance metrics of robotic suturing and knot tying of an anastomotic model were developed. They were then presented to very experienced surgeons in a modified Delphi process. We achieved consensus on objectively defined performance metrics (i.e., procedure steps, errors, and critical errors) thus demonstrating face and content validity of a basic anastomotic suturing and knot tying task. We then compared the performance of experienced robotic surgeons and (trained) novices performing the ORSI suturing and knot tying chicken model. The metrics statistically significantly distinguished between the performance of the two groups.

Compliance with ethical standards

Disclosures Stefano Puliatti, Elio Mazzone, Marco Amato, Ruben De Groote, Alexandre Mottrie, and Anthony G. Gallagher declare that they have no conflicts of interest or financial ties to disclose.

References

1. George EI, Brand TC, LaPorta A, Marescaux J, Satava RM (2018) Origins of robotic surgery: from skepticism to standard of care. JSLS. https://doi.org/10.4293/JSLS.2018.00039
2. Palagonia E, Mazzone E, De Naeyer G, D'Hondt F, Collins J, Wisz P, Van Leeuwen FWB, Van Der Poel H, Schatteman P, Mottrie A, Dell'Oglio P (2019) The safety of urologic robotic surgery depends on the skills of the surgeon. World J Urol 38:1373–1383
3. Puliatti S, Mazzone E, Dell'Oglio P (2020) Training in robot-assisted surgery. Curr Opin Urol 30:65–72
4. Pierorazio PM, Allaf ME (2009) Minimally invasive surgical training: challenges and solutions. Urol Oncol 27:208–213
5. Alemzadeh H, Raman J, Leveson N, Kalbarczyk Z, Iyer RK (2016) Adverse events in robotic surgery: a retrospective study of 14 years of FDA Data. PLoS ONE 11:e0151470
6. Angelo RL, Ryu RK, Pedowitz RA, Beach W, Burns J, Dodds J, Field L, Getelman M, Hobgood R, McIntyre L, Gallagher AG (2015) A proficiency-based progression training curriculum coupled with a model simulator results in the acquisition of a superior arthroscopic bankart skill set. Arthroscopy 31:1854–1871
7. Breen D, O'Brien S, McCarthy N, Gallagher A, Walshe N (2019) Effect of a proficiency-based progression simulation programme on clinical communication for the deteriorating patient: a randomised controlled trial. BMJ Open 9:e025992
8. Kallidaikurichi Srinivasan K, Gallagher A, O'Brien N, Sudir V, Barrett N, O'Connor R, Holt F, Lee P, O'Donnell B, Shorten G (2018) Proficiency-based progression training: an 'end to end' model for decreasing error applied to achievement of effective epidural analgesia during labour: a randomised control study. BMJ Open 8:e020099
9. Cates CULL, Gallagher AG (2016) Prospective, randomised and blinded comparison of proficiency-based progression full-physics virtual reality simulator training versus invasive vascular experience for learning carotid artery angiography by very experienced operators. BMJ Simul Technol Enhanc Learn 2(1):1–5
10. Pedowitz RA, Nicandri GT, Angelo RL, Ryu RK, Gallagher AG (2015) Objective assessment of knot-tying proficiency with the fundamentals of arthroscopic surgery training program workstation and knot tester. Arthroscopy 31:1872–1879
11. Van Sickle KR, Ritter EM, Baghai M, Goldenberg AE, Huang IP, Gallagher AG, Smith CD (2008) Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. J Am Coll Surg 207:560–568
12. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA 3rd, Ramel S, Smith CD, Arvidsson D (2007) Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. Am J Surg 193:797–804
13. Vanlander AE, Mazzone E, Collins JW, Mottrie AM, Rogiers XM, van der Poel HG, Van Herzeele I, Satava RM, Gallagher AG (2020) Orsi Consensus Meeting on European Robotic Training (OCERT): results from the First Multispecialty Consensus Meeting on Training in Robot-assisted Surgery. Eur Urol. https://doi.org/10.1016/j.eururo.2020.02.003
14. Cacciamani G, De Marco V, Siracusano S, De Marchi D, Bizzotto L, Cerruto MA, Motton G, Porcaro AB, Artibani W (2017) A new training model for robot-assisted urethrovesical anastomosis and posterior muscle-fascial reconstruction: the Verona training technique. J Robot Surg 11:123–128
15. Santos DRD, Calvo FC, Feijo DH, Araujo NP, Teixeira RKC, Yasojima EY (2019) New training model using chickens intestine for pediatric intestinal anastomosis. Acta Cir Bras 34:e201900709
16. Sotelo RJ, Astigueta JC, Carmona OJ, De Andrade RJ, Moreira OE (2009) Chicken gizzard: a new training model for laparoscopic urethrovesical anastomosis. Actas Urol Esp 33:1083–1087
17. Volpe A, Ahmed K, Dasgupta P, Ficarra V, Novara G, van der Poel H, Mottrie A (2015) Pilot validation study of the European Association of Urology robotic training curriculum. Eur Urol 68:292–299
18. Van Velthoven RF, Ahlering TE, Peltier A, Skarecky DW, Clayman RV (2003) Technique for laparoscopic running urethrovesical anastomosis:the single knot method. Urology 61:699–702
19. Gallagher AG (2012) Metric-based simulation training to proficiency in medical education: what it is and how to do it. Ulster Med J 81:107–113
20. Gallagher AG, O'Sullivan GC (2011) Fundamentals of surgical simulation: principles and practice. Springer Science & Business Media, London
21. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, Smith CD, Satava RM (2005) Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. Ann Surg 241:364–372
22. Kojima K, Graves M, Taha W, Cunningham M, Joeris A, Gallagher AG (2018) AO international consensus panel for metrics on a closed reduction and fixation of a 31A2 pertrochanteric fracture. Injury 49:2227–2233
23. Gallagher AG, Ryu RKN, Pedowitz RA, Henn P, Angelo RL (2018) Inter-rater reliability for metrics scored in a binary fashion-performance assessment for an arthroscopic bankart repair. Arthroscopy 34:2191–2198
24. Kazdin AE (1994) Behavior modification in applied settings. Brooks. Cole Publishing Co, Pacific Grove, CA

25. Mazzone E, Mistretta FA, Knipper S, Tian Z, Larcher A, Widmer H, Zorn K, Capitanio U, Graefen M, Montorsi F, Shariat SF, Saad F, Briganti A, Karakiewicz PI (2019) Contemporary national assessment of robot-assisted surgery rates and total hospital charges for major surgical uro-oncological procedures in the United States. J Endourol 33:438–447
26. Brook NR, Dell'Oglio P, Barod R, Collins J, Mottrie A (2019) Comprehensive training in robotic surgery. Curr Opin Urol 29:1–9
27. Larcher A, De Naeyer G, Turri F, Dell'Oglio P, Capitanio U, Collins JW, Wiklund P, Van Der Poel H, Montorsi F, Mottrie A (2019) The ERUS curriculum for robot-assisted partial nephrectomy: structure definition and pilot clinical validation. Eur Urol 75:1023–1031
28. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM (2002) Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 236:458–463 **discussion 463–454**
29. Breen D, O'Brien S, McCarthy N, Gallagher A, Walshe N (2019) Effect of a proficiency-based progression simulation programme on clinical communication for the deteriorating patient: a randomised controlled trial. BMJ Open 9:e025992
30. Srinivasan KK, Gallagher A, O'Brien N, Sudir V, Barrett N, O'Connor R, Holt F, Lee P, O'Donnell B, Shorten G (2018) Proficiency-based progression training: an 'end to end' model for decreasing error applied to achievement of effective epidural analgesia during labour: a randomised control study. BMJ Open 8:e020099
31. Cates CU, Lönn L, Gallagher AG (2016) Prospective, randomised and blinded comparison of proficiency-based progression full-physics virtual reality simulator training versus invasive vascular experience for learning carotid artery angiography by very experienced operators. BMJ Simul Technol Enhanced Learn 2:1–5
32. Angelo RL, Ryu RK, Pedowitz RA, Beach W, Burns J, Dodds J, Field L, Getelman M, Hobgood R, McIntyre L, Gallagher AG (2015) A proficiency-based progression training curriculum coupled with a model simulator results in the acquisition of a superior arthroscopic Bankart skill set. Arthrosc: J Arthrosc Relat Surg 31:1854–1871