Proficiency-based progression training for robotic surgery skills training

De Groote, R., Puliatti, S., Amato, M., Mazzone, E., Rossiello, G., Farinha, R., Paludo, A., Desender, L., Van Cleynenbreugel, B., Bunting, B. P., Mottrie, A., Gallagher, A. G., Larcher, A., Uvin, P., Decoene, J., Tuyten, T., D'Hondt, M., Hubert, N., Chatzopoulos, C., & De Troyer, B. (2022). Proficiency-based progression training for robotic surgery skills training: a randomized clinical trial. *BJU International*, *130*(4), 528-535. https://doi.org/10.1111/bju.15811

Published in:
BJU International

Publication Status:
Published (in print/issue): 31/10/2022

DOI:
10.1111/bju.15811

Document Version
Author Accepted version

De Groote ruben (Orcid ID: 0000-0003-4081-8221)
Mazzone Elio (Orcid ID: 0000-0003-1932-1800)
Paludo Artur de Oliveira (Orcid ID: 0000-0002-0701-3301)
Van Cleynenbreugel Ben S.E.P. (Orcid ID: 0000-0001-9517-3673)

Proficiency based progression training for robotic skill training: a randomized clinical trial

On behalf of the Junior ERUS / YAU working group on robot-assisted surgery of the European Association of Urology and the ERUS Education Working Group.

Authors:
Ruben De Groote, MD[1,2], Stefano Puliatti, MD[1,2,3], Marco Amato, MD[1,2,3], Elio Mazzone, MD[4], Giuseppe Rossiello, MD[1,2,4], Rui Farinha, MD[1,2], Artur Paludo, MD[1,2,5], Liesbeth Desender, MD, PhD[6], Ben Van Cleynenbreugel, MD, PhD[7], Brendan P. Bunting, PhD[8], Alexandre Mottrie, MD, PhD[1,2] and Anthony G. Gallagher, PhD, DSc[1,7,9]

Collaborators:
Alessandro Larcher, MD[4], Pieter Uvin, MD, PhD[10], Jasper Decoene, MD[11], Tom Tuyten, MD[12], Mathieu D'Hondt, MD[13], Nicolas Hubert, MD[14], Charles Chatzopoulos, MD[15], Bart De Troyer, MD[16]

[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; Vita-Salute San Raffaele University, Milan, Italy

[5] Clinic Hospital of Porto Alegre, Urology, Porto Alegre, Brazil

[6] Department of Thoracovascular Surgery, University Hospital Ghent, Ghent, Belgium

[7] Faculty of Medicine, KU Leuven, Leuven, Belgium

[8] School of Psychology, Ulster University, Coleraine, Northern Ireland, UK

[9] School of Medicine, Faculty of Life and Health Sciences, Ulster University, Northern Ireland, UK.

[10] Department of Urology, AZ Sint-Jan, Bruges, Belgium

[11] Department of Urology, OLV van Lourdes Hospital, Waregem, Belgium

[12] Department of Urology, Jessa Hospital, Hasselt, Belgium

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/bju.15811

[13] Department of Surgery, AZ Groeninge, Kortrijk, Belgium

[14] Department of Urology, CHR de la Citadelle, Liège, Belgium

[15] Department of Urology, Chirec Hospital, Brussels, Belgium

[16] Department of Urology, AZ Nikolaas, Sint-Niklaas, Belgium

EAU Robotic Urology Section (ERUS) Scientific Committee sponsorship of the 7 Belgian surgeons should be acknowledged.

Corresponding author:

Alexandre Mottrie, MD, PhD

Consultant Urological Surgeon, Uro-oncology and Robotics
Department of Urology
Onze Lieve Vrouw Ziekenhuis Aalst
Moorselbaan 164, 9300 Aalst
Belgium
E-mail: a.mottrie@gmail.com
Telephone: +32475/225250
Fax: +3253724411

Proficiency based progression training for robotic skill training: a randomized clinical trial

On behalf of the Junior ERUS / YAU working group on robot-assisted surgery of the European Association of Urology and the ERUS Education Working Group.

Authors:
Ruben De Groote, MD[1,2], Stefano Puliatti, MD[1,2,3], Marco Amato, MD[1,2,3], Elio Mazzone, MD[4], Giuseppe Rossiello, MD[1,2,4], Rui Farinha, MD[1,2], Artur Paludo, MD[1,2,5], Liesbeth

Desender, MD, PhD[6], Ben Van Cleynenbreugel, MD, PhD[7], Brendan P. Bunting, PhD[8], Alexandre Mottrie, MD, PhD[1,2] and Anthony G. Gallagher, PhD, DSc[1,7,9]

Collaborators:
Alessandro Larcher, MD[4], Pieter Uvin, MD, PhD[10], Jasper Decoene, MD[11], Tom Tuyten, MD[12], Mathieu D'Hondt, MD[13], Nicolas Hubert, MD[14], Charles Chatzopoulos, MD[15], Bart De Troyer, MD[16]

[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;

Vita-Salute San Raffaele University, Milan, Italy

[5] Clinic Hospital of Porto Alegre, Urology, Porto Alegre, Brazil

[6] Department of Thoracovascular Surgery, University Hospital Ghent, Ghent, Belgium

[7] Faculty of Medicine, KU Leuven, Leuven, Belgium

[8] School of Psychology, Ulster University, Coleraine, Northern Ireland, UK

[9] School of Medicine, Faculty of Life and Health Sciences, Ulster University, Northern Ireland, UK.

[10] Department of Urology, AZ Sint-Jan, Bruges, Belgium

[11] Department of Urology, OLV van Lourdes Hospital, Waregem, Belgium

[12] Department of Urology, Jessa Hospital, Hasselt, Belgium

[13] Department of Surgery, AZ Groeninge, Kortrijk, Belgium

[14] Department of Urology, CHR de la Citadelle, Liège, Belgium

[15] Department of Urology, Chirec Hospital, Brussels, Belgium

[16] Department of Urology, AZ Nikolaas, Sint-Niklaas, Belgium

EAU Robotic Urology Section (ERUS) Scientific Committee sponsorship of the 7 Belgian surgeons should be acknowledged.

Corresponding author:

Alexandre Mottrie, MD, PhD

Consultant Urological Surgeon, Uro-oncology and Robotics
Department of Urology
Onze Lieve Vrouw Ziekenhuis Aalst
Moorselbaan 164, 9300 Aalst
Belgium
E-mail: a.mottrie@gmail.com
Telephone: +32475/225250
Fax: +3253724411

Running head: PBP vs Traditional training: an RCT

ABBREVIATIONS

PBP	Proficiency Based Progression
RCT	Randomized Controlled Trial
VUA	Vesico-Urethral Anastomosis
RARP	Robot-Assisted Radical Prostatectomy
TTG	Traditional Trained Group
ToT	Transfer of Training
IRR	Inter-Rater Reliability
CI	Confidence Interval
ISBAR	Identification, Situation, Background, Assessment, Recommendation
ERUS	EAU Robotic Urology Section

ABSTRACT

Objective To determine whether Proficiency Based Progression (PBP) training leads to better robotic surgical performance compared to traditional training as the value of PBP training for learning robotic surgical skills is unclear.

Material and Methods The PROVESA trial is a multicentric, prospective, randomized and blinded clinical study comparing PBP training with traditional training for robotic suturing and knot tying anastomosis skills. 36 Robotic naïve junior residents were recruited from 16 training sites and 12 residency training programs. Participants were randomly allocated to metric-based PBP training or the current standard of care Traditional Training, and compared at the end of training. Primary outcome was percentage of subjects reaching the predefined proficiency benchmark. Secondary outcome was the number of procedure steps and errors made.

Results: 17% of the Traditional Trained Group and 67% of the PBP group demonstrated the proficiency benchmark (i.e., ~ 10 times as likely to demonstrate proficiency (p=0.006)). The PBP group demonstrated a 51% reduction in performance errors from baseline to the final assessment (18.3 vs 8.9). The Traditional Trained Group demonstrated marginal improvement (15.94 vs 15.44) in errors made.

Conclusions: The PROVESA trial is the first prospective RCT for basic skill training in robotic surgery. Implementation of a PBP training methodology resulted in superior surgical performance for robotic suturing and knot tying anastomosis performance. Compared to traditional training, a better surgical quality could be obtained by implementing PBP for robotic basic skill training.

KEYWORDS

Basic skill training, Proficiency based progression, Robotic surgery, Training, Surgical simulation

INTRODUCTION

Halsted's apprenticeship model is still widely used during the training of surgeons (1,2). Novice surgical residents still perform their first surgical steps on patients, risking higher peri-operative morbidity (3,4). Surgical training depends on quality of the trainer with a lack of standardization, objective, transparent and fair feedback(5). Reduced work hours, increased bureaucracy and ethical concerns have forced the surgical community to explore new training modalities. In similarity with training of airline pilots, mitigating the initial learning curve in an out-of-hospital setting, and developing a standardized, simulation-based training system, could improve patient safety (6).

Proficiency Based Progression (PBP) has demonstrated its value in different surgical specialties (7,8). Cornerstone of PBP training are procedure-specific, validated and binary performance metrics derived from characterization of this procedure done by experts with well-defined procedural steps and errors. A specific level of training outcome, defined by a quantitative score (benchmark) on a standardized assessment, must be demonstrated to gain the proficiency level. Progression in training is only allowed by meeting the quantitatively defined performance benchmarks (7). In a meta-analysis, an average reduction of metrics-based performance errors of 60% was observed in multiple specialties compared to traditional training(9).

Fundamental to creating a proficient robotic surgeon is the acquisition of basic surgical skills(10). The "Venezuelan chicken model" is generally accepted as an optimal model to learn robotic suturing, anastomosis and knot tying skills. It is, for example, a good model for the vesico-urethral anastomosis (VUA) training part of a robot-assisted radical prostatectomy (RARP). The development of performance metrics and its construct validity have been reported by Puliatti et al (11). This randomized controlled trial (RCT) aims to compare the effectiveness of the apprenticeship approach to training with the PBP approach for teaching robotic suturing and knot tying of a VUA on a chicken model for novice surgical residents.

MATERIAL AND METHODS

Participants/Subjects

The study subjects were robotic naïve, first and second year surgical residents (n=36) from Ghent University and KU Leuven. All included participants were indeed robotic naïve, meaning that they did not have any previous experience in suturing and knot tying using robotic technology. Any experience on the robotic console as operator was deemed to be an exclusion criterium. Approval for the study was received from the local ethical committee and the trial was registered at the National Institution of Health (NCT04786834). All subjects completed informed consents.

Faculty Training

The PROVESA PBP faculty were supervised by the PROVESA research coordinator, a consultant behavioral scientist (A.G.G.). The panel members were randomly assigned to four fixed pairs which remained constant through scoring of all live and video-recorded performances. Prior to commencement of the trial raters were trained to mastery of the metrics which was indicated by repeated inter-rater reliability > 0.8 between pairs of reviewers for objective assessment of full-length surgical videos.

In the apprenticeship model, a resident is depending, for its surgical exposure, on its mentor and the institution where he or she is trained. Consequently, when recruiting first and second year surgical residents, participants with different levels of surgical exposure are gathered. Therefore, to rule out selection bias based on surgical skills, all participants had a baseline assessment of their first VUA. Then, a matched 1:1 randomization was performed (using an online randomizer www.random.org). Subjects were matched for age (± 2 years), residency year and skill at baseline as determined by the objectively assessed performance metrics score (Figure 1a & b). The Traditional Trained Group (TTG) were yoked to the same training time as their counterpart in the PBP Group.

Group PBP training:

Eighteen participants were randomized to the PBP group. They were given access to a dedicated e-course on the online Bridge® platform one week before their training in the skills lab of ORSI Academy (Figure 1a & b). In this e-course, the validated operative metrics were reviewed during which errors and steps were illustrated (Appendix 1). Immediate performance feedback was given to the participants during the course. Training in the skills lab required all subjects to reach the quantitively defined proficiency benchmark (94%) on the eLearning module (defined as the mean score on the test by the panel of experienced surgeons). Training and assessment methods, including the VUA task are described in more detail in Supplementary information (appendix 2).

In the skills lab training was delivered to teams of three participants per trainer during a full day by the PROVESA PBP faculty in a standardized way with the operative metrics as guiding instrument. While one subject was training, other two participants scored and gave metric-base formative feedback to their colleague on task completion. Each group of three trainees was supervised by an experienced faculty member who gave ongoing metric-based performance feedback (i.e., deliberate practice training)(12). Participants could train as long as they liked, and faculty indicated when they thought the trainee had reached the proficiency level. Subsequently, the final trial was supervised, and videorecorded by one of the designated PROVESA faculty members.

Group Traditional Training:

Eighteen participants were randomly assigned to the TTG. Surgical trainers from different, recognized robotic training centers were invited to train the TTG based on the way surgical residents currently get trained. For one week, trainees had continuous access to the exact same online learning platform as the PBP Group and were repeatedly encouraged to study the material._Subsequently, they were invited to ORSI Academy for a full day of training by seven robotic experts. All experts were selected from different Belgian hospitals and were considered to be experts in robotic surgery and excellent trainers. Each trainer had three participants to train. As matching was done for training time with their counterparts in the PBP group, every participant in the TTG had a preset number of trials before doing the final assessment.

Video Scoring

The PROVESA research coordinator randomly assigned the 36 full-length study videos, each with only the designated unique identifying number attached to a single pair of reviewers. Other than the research coordinator, all video reviewers remained blinded to the subject and training condition of the video being reviewed. Each video was independently reviewed and scored by the 2 members of an assigned pair of reviewers. Performance was scored in a binary fashion for performance units in the operative metrics that were or were not observed to have occurred by the reviewer.

Statistical Analysis

Statistical Analysis was performed with SPSS 26 (Armonk, New York). Differences between the relationship of the two training programs on proficiency demonstration (the primary end point) at the end of training were examined using logistic regression analysis. A 2 * 2 Mixed

Model analysis of variance (ANOVA) was used to determine if there was a statistical difference for secondary end points (the number of completed steps and the total number of errors made) between the Traditional Trained and the PBP trained groups 1) at baseline and 2) post training assessments (i.e., repeated measures) with specific within-subject contrasts for Steps and Errors compared with t-distribution in the mixed model ANOVA. Results are reported in terms of the statistical estimate (Est), the standard errors (SE), degree of freedom (df), test statistic (t) and the probability value (p).

Design

Data at time one relates to Baseline, while that at time two refers to the post training performance level, and within group comparisons where individuals act as their own control. Respondents were randomly assigned to one of two conditions (a) Traditionally Trained and (b) Proficiency Based Progression Groups. Of particular interest in this design is the cross-level interaction of assessment occasion (i.e., Baseline Vs Post training) by Condition (i.e., TTG vs PBP), where it is anticipated that PBP training should perform better.

Statistical Power Calculations

Power calculation: the numbers needed in each arm were based on transfer of training (ToT) effects observed in previous studies of PBP simulation studies where ToT rates of 42-69% were observed (5,13–19). In the current study we therefore expected to observe differences of 6.9% v 60%(13) and 29% vs 75%(17) which generated statistical power of 0.96 and 0.82 respectively for proficiency demonstration of 18 trainees in each group at the end of training. Likewise, a decrease in performance errors >40%, with a two-tailed test, and with n = 4 trainees in each group with an alpha of 5% (which corresponds to a 95% confidence interval) and Beta error 10% (i.e., 1-0.1=0.9 βeta) would yield a statistical power of 95%.

RESULTS

Inter-rater reliability

Baseline characteristics with respect to age, gender, handedness, surgical discipline etc., of the participants in each group are shown in Table 1. The mean IRR of video recorded performance assessments for all metrics was 0.87 (TTG IRR=0.86 and PBP Trained Group=0.87). None of the video recorded assessments were below the 0.8 IRR level.

A significant difference between the two groups eLearning scores (PBP=96.9 (2) vs TTG=83 (7), $F=29.02$, $p<0.000$) was observed.

Descriptive statistics

Figure 2 shows percentages of participants in each group who completed all five procedure steps at the baseline and post training assessments. At the baseline assessment 57% of the TTG and 43% of the PBP Group completed all five steps of the procedure. After training this had increased for both groups; 89% (n=16/18) in the Traditional and 94% (n=17/18) for the PBP Group. The post training difference between the two groups was not statistically significant (Pearson Chi^2=0.364, $p=0.5$).

Logistic regression

The post training assessment showed that 17% of the TTG trainees (n=3/18) and 67% (n=12/18) of the PBP trained Group demonstrated the proficiency benchmark. In a logistic regression analysis (Figure 3c) it was found that in comparison to the TTG, the PBP group were >10 times as likely to demonstrate proficiency which was statistically different (Table 2 - Exp (B)=10.00 (95% CI of Exp (B), Lower=0.722 - Upper=3.88), Wald Chi-Squared=8.157, $p=0.004$).

On the post training assessment 39% (n=7/18) of the TTG had an anastomosis leak or failed to complete the procedure in comparison to the PBP Trained group which had a rate of 17.6% (n=3/17). Although the TTG had more than twice the rate of the PBP group, this difference was not statistically significant (Pearson $Chi^2=1.93$, p=0.164).

Multivariate statistics

Procedure Steps

The overall mean and 95% CI for the number of procedure Steps completed by both groups at Baseline and Post training assessments are shown in Figure 3a. The TTG showed a 66% within group performance improvement as well as a reduction in performance variability at post training assessment. At baseline, on average those in the TTG condition completed 2.9 procedure Steps which had increased to 4.8 in the post training assessment (t=-3.929, p=0.001). A similar pattern was observed for the PBP trained group in terms of improved performance and reduced performance variability. At baseline, on average they completed 2.7 procedure Steps which had increased to 4.9 in the post training assessment (t=-4.955, p<0.000). As a main effect for both groups, this difference was statistically significant for Baseline to Post Training Assessment (95% CI of the difference, Lower=-2.97 – Upper=-0.92, df=37.76, t=-3.831, p <0.000). No difference between the two Groups nor Group * Repeated Measure interaction was observed (Group, t=0.338, p=0.738 and for Group * Repeated Measure, t=-0.31, p=0.759)

Errors

The same analysis was completed for the number of errors made by each group (Figure 3b). The TTG demonstrated only a marginal performance improvement from Baseline to the Post training assessment (15.94 vs 15.44). In contrast the PBP trained group demonstrated a 51% reduction in performance errors from Baseline to the Post training assessment (18.3 vs 8.9).

The overall difference between Groups was found to be statistically significant (Table 2: est=-6.67, se=1.62, df=34, t=-4.11, p<0.000)

The effect of test occasions on the number of errors reported/committed was different depending on the group allocation. Those in the PBP condition made fewer errors (est=9.06, se=1.62, df=62.06, t=3.282, sig=0.002) during the Post Training assessment.

DISCUSSION

Suturing and knot tying are fundamental skills in the training of robotic surgeons. Our study represents the first RCT which investigated the possible added value of PBP in the training of novice robotic surgeons. After 1 day of training 67% of the PBP participants reached the quantitatively defined proficiency benchmark as compared to 17% in the traditional group. Consequently, PBP subjects were 10 times more likely to achieve proficiency. Had more time been available for training the number of trainees demonstrating proficiency would have been even greater (i.e., 83 – 89%)(20).

No significant difference was observed between the number of procedural steps completed by the study groups, at baseline and in their final assessment. Conversely, there was a considerable reduction in procedural errors made during the final assessment in the PBP group indicating the key difference between both training methodologies. Where traditional training emphasizes what to do i.e., procedural steps, PBP training also emphasizes what not to do, i.e., performance errors because completion of the procedure does not guarantee the quality of the surgical performance.

The valid and reliable assessment of surgical skills is imperative to effective and efficient training. The Institute of Medicine has argued convincingly that medicine must move away from a process driven approach (e.g., time in training, number of procedures done etc.) (21) for the evaluation of skills to an outcome based approach, where the skills of the trainee are verified. Furthermore, evidence has now started to emerge which demonstrates a strong relationship between operative skills and patient outcomes, i.e., morbidity and mortality. Birkmeyer et al., have shown that better surgical skills lead to better outcomes (3). Consequently, the importance of surgical training to develop skill amongst surgeons cannot be underestimated. Conversely, it is shown by George et al, that 33.3% of US General Surgery residents are not ready to independently perform core procedures by the time they complete residency training (22). A substantial increase of procedural types, with higher complexity and skill requirement, combined with restriction on working hours for residents will further stress the current training paradigm.

PBP training is a novel training methodology that approaches the airline pilot training model with the incorporation of surgical simulation and validated, binary performance metrics, and standardized procedure templates as cornerstones of training (23,24). Angelo et al., has shown in a prospective, blinded RCT that implementation of a PBP curriculum coupled with the use of a shoulder model simulator produced a superior arthroscopic surgical skill-set when compared with traditional and simulator-enhanced training methods (17,20,25). In a RCT by Breen et al., it was also shown that implementing PBP for the training of the ISBAR (Identification, Situation, Background, Assessment, Recommendation) communication skillset in the context of a clinically deteriorating patient led to a significantly better communication performance(13). In a systematic review, Mazzone et al. reported that the implementation of a PBP curriculum led to an average reduction 60% in performance errors compared to traditional training methodologies(9). A logical conclusion from these studies and the results reported here is that PBP is a more effective approach to skills training.

Study Limitations

Inevitably, this study was associated with some limitations. Although very well controlled, this was a small scale study, powered mainly for difference in proficiency level and performance errors(8,13,17). Therefore, it was not possible to show a significant difference in VUA leakage, although this was double the observed rate in the Traditional Group versus the PBP group. Larger future studies will be required to elucidate whether this reduction of performance errors could be translated into a reduction in adverse events. Nevertheless, as the first RCT in robotic surgery indicating a superior outcome of PBP training versus traditional training for basic skill training, the value of this paper is of high importance.

CONCLUSION

The PROVESA trial has shown that implementation of PBP methodology, compared to traditional training, resulted in superior surgical performance for robotic suturing and knot tying. Large RCT's are necessary to confirm these results and further elucidate the clinical benefit of PBP methodology.

CONFLICTS OF INTEREST

None declared.

ACKNOWLEDGMENTS

This study was conducted on behalf of the Junior ERUS / YAU working group on robot-assisted surgery of the European Association of Urology and the ERUS Education Working Group. EAU Robotic Urology Section (ERUS) Scientific Committee sponsorship of the 7 Belgian surgeons should be acknowledged.

REFERENCES

1. Pellegrini CA. Surgical education in the United States: Navigating the white waters. In: Annals of Surgery [Internet]. Ann Surg; 2006 [cited 2021 Jan 3]. p. 335–42. Available from: https://pubmed.ncbi.nlm.nih.gov/16926559/
2. Chikwe J, De Souza AC, Pepper JR. No time to train the surgeons [Internet]. Vol. 328, British Medical Journal. BMJ Publishing Group; 2004 [cited 2020 Nov 11]. p. 418–9. Available from: https://pubmed.ncbi.nlm.nih.gov/14976074/
3. Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR, et al. Surgical Skill and Complication Rates after Bariatric Surgery. N Engl J Med. 2013 Oct 10;369(15):1434–42.
4. Palagonia E, Mazzone E, De Naeyer G, D'Hondt F, Collins J, Wisz P, et al. The safety of urologic robotic surgery depends on the skills of the surgeon. World J Urol [Internet]. 2020 Jun 1 [cited 2020 Nov 8];38(6):1373–83. Available from: https://pubmed.ncbi.nlm.nih.gov/31428847/
5. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance results of a randomized, double-blinded study. In: Annals of Surgery. 2002. p. 458–64.
6. Logan M. Why Hospitals Should Fly: The Ultimate Flight Plan to Patient Safety and Quality Care. Biomed Instrum Technol. 2010 Mar 1;44(2):110–110.
7. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. Vol. 241, Annals of Surgery. 2005. p. 364–72.
8. Gallagher AG. Metric-based simulation training to proficiency in medical education:- What it is and how to do it. Vol. 81, Ulster Medical Journal. Ulster Medical Society; 2012. p. 107–13.
9. Mazzone Elio PSAMBBRBMFMAGA. A Systematic Review and Meta-Analysis on the Impact of Proficiency-Based Progression Simulation Training on Performance Outcomes. Ann Surg [Internet]. 2020 [cited 2021 Jun 13]; Available from: https://journals.lww.com/annalsofsurgery/Abstract/9000/A_Systematic_Review_and_Meta_Analysis_on_the.93826.aspx
10. Vanlander AE, Mazzone E, Collins JW, Mottrie AM, Rogiers XM, van der Poel HG, et al. Orsi Consensus Meeting on European Robotic Training (OCERT): Results from the First Multispecialty Consensus Meeting on Training in Robot-assisted Surgery. Vol. 78, European Urology. Elsevier B.V.; 2020. p. 713–6.
11. Puliatti S, Mazzone E, Amato M, De Groote R, Mottrie A, Gallagher AG. Development and validation of the objective assessment of robotic suturing and knot tying skills for chicken anastomotic model. Surg Endosc [Internet]. 2020 [cited 2021 Jan 3]; Available from: https://pubmed.ncbi.nlm.nih.gov/32857241/
12. Ericsson KA, Krampe RT, Tesch-Römer C. The Role of Deliberate Practice in the Acquisition of Expert Performance. Psychol Rev [Internet]. 1993 [cited 2021 Feb 18];100(3):363–406. Available from: /record/1993-40718-001
13. Breen D, O'Brien S, McCarthy N, Gallagher A, Walshe N. Effect of a proficiency-based progression simulation programme on clinical communication for the deteriorating patient: A randomised controlled trial. BMJ Open [Internet]. 2019 Jul 1 [cited 2020 Nov 8];9(7). Available from: https://pubmed.ncbi.nlm.nih.gov/31289064/
14. Kallidaikurichi Srinivasan K, Gallagher A, O'Brien N, Sudir V, Barrett N, O'Connor R, et al. 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 [Internet]. 2018 Oct 1 [cited 2021 Jan 3];8(10).

Available from: https://pubmed.ncbi.nlm.nih.gov/30327396/
15. Cates CU, Lönn L, Gallagher AG. 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 [Internet]. Vol. 2, BMJ Simulation and Technology Enhanced Learning. BMJ Publishing Group; 2016 [cited 2021 Jan 3]. p. 1–5. Available from: http://stel.bmj.com/
16. Pedowitz RA, Nicandri GT, Angelo RL, Ryu RKN, Gallagher AG. Objective Assessment of Knot-Tying Proficiency with the Fundamentals of Arthroscopic Surgery Training Program Workstation and Knot Tester. Arthrosc - J Arthrosc Relat Surg [Internet]. 2015 Oct 1 [cited 2021 Jan 3];31(10):1872–9. Available from: https://pubmed.ncbi.nlm.nih.gov/26298642/
17. Angelo RL, Ryu RKN, Pedowitz RA, Beach W, Burns J, Dodds J, et al. 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. 2015 Oct 1;31(10):1854–71.
18. Van Sickle KR, Ritter EM, Baghai M, Goldenberg AE, Huang IP, Gallagher AG, et al. Prospective, Randomized, Double-Blind Trial of Curriculum-Based Training for Intracorporeal Suturing and Knot Tying. J Am Coll Surg. 2008 Oct;207(4):560–8.
19. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. Am J Surg [Internet]. 2007 Jun [cited 2021 Jan 3];193(6):797–804. Available from: https://pubmed.ncbi.nlm.nih.gov/17512301/
20. Angelo RL, Tauro J, St. Pierre P, Ross G, Voloshin I, Shafer B, et al. Arthroscopic Rotator Cuff Repair Metrics: Establishing Face, Content, and Construct Validity in a Cadaveric Model. Arthrosc - J Arthrosc Relat Surg. 2020 Jan 1;36(1):71-79.e1.
21. Asch DA, Weinstein DF. Innovation in Medical Education. N Engl J Med [Internet]. 2014 Aug 28 [cited 2021 Feb 8];371(9):794–5. Available from: https://pubmed.ncbi.nlm.nih.gov/25075620/
22. George BC, Bohnen JD, Williams RG, Meyerson SL, Schuller MC, Clark MJ, et al. Readiness of US General Surgery Residents for Independent Practice. Ann Surg [Internet]. 2017 Oct 1 [cited 2020 Nov 11];266(4):582–94. Available from: https://pubmed.ncbi.nlm.nih.gov/28742711/
23. Gallagher AG, O'Sullivan GC, Gallagher AG, O'Sullivan GC, O'Sullivan GC. Simulations for Procedural Training. In: Fundamentals of Surgical Simulation. Springer London; 2011. p. 39–66.
24. Fried MP, Gallagher AG, Satava RM. Training to proficiency: Aircraft to OR [Internet]. Vol. 130, Archives of Otolaryngology - Head and Neck Surgery. Arch Otolaryngol Head Neck Surg; 2004 [cited 2021 Jan 10]. p. 1145–6. Available from: https://pubmed.ncbi.nlm.nih.gov/15492158/
25. Angelo RL, Pedowitz RA, Ryu RKN, Gallagher AG. The Bankart Performance Metrics Combined with a Shoulder Model Simulator Create a Precise and Accurate Training Tool for Measuring Surgeon Skill. Arthrosc - J Arthrosc Relat Surg. 2015 Sep 1;31(9):1639–54.

FIGURE LEGENDS

Figure 1a and b. a) The design of the multicenter, prospective, randomized, matched and blinded study and b) The CONSORT 2010 flow diagram of the PROVESA trial.

Figure 2. The percentage of trainees who completed all five procedure steps at baseline assessment and at the end of their training in the video-recorded and blinded assessment.

Figure 3a - c. The mean and 95% confidence intervals (CI) of a.) procedure steps completed and b.) the number of errors made by the Traditional and PBP trained Groups at Baseline and at the end of training and c.) Logistic Regression Analysis for the relative differences between the Traditional Trained Group (Reference Group for comparison) and the PBP trained group proficiency demonstration at the end of training.

	Traditional Trained (n=18)	**PBP Trained (n=18)**	prob. level
Age	25 (1)	26 (2)	0.323
Gender	12F	8F	0.314
Handed	13R	15R	0.437
PG Year	1.2 (1)	1.2 (1)	1.000

() = standard deviation; PG = Postgraduate; F = female; R = Right-handed

Table 1. Demographic information for Traditional Trained group and PBP group

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
	Number of Errors						
Intercept	15.44	1.147	34	13.470	.000	13.114	17.775
Group=1.00	-6.67	1.622	34	-4.111	.000	-9.962	-3.372
Assessment Occasion=1.00	.50	1.951	62.065	.256	.799	-3.4	4.4

Assessment Occasion=1.00 * Group=1.00	9.06	2.759	62.065	3.282	.002	3.54	14.571
	Number of Steps						
Intercept	4.83	.116	34	41.559	.000	4.597	5.07
Group=1	.06	.164	34	.338	.738	-.279	.39
Assessment Ocassion	-1.94	.508	37.756	-3.831	.000	-2.972	-.917
Assessment Occasion * Group=1.00	-.22	.718	37.756	-.310	.759	-1.676	1.231

Group=1: PBP, Group=2: Control (value set at zero and omitted)
Assessment Occasion=1: Baseline, Occasion=2: Time two (value set at zero and omitted)
All redundant parameters in the interaction term have been omitted.

Table 2. Estimates of a 2x2 fixed effects model representing the number of errors made and the number of steps taken

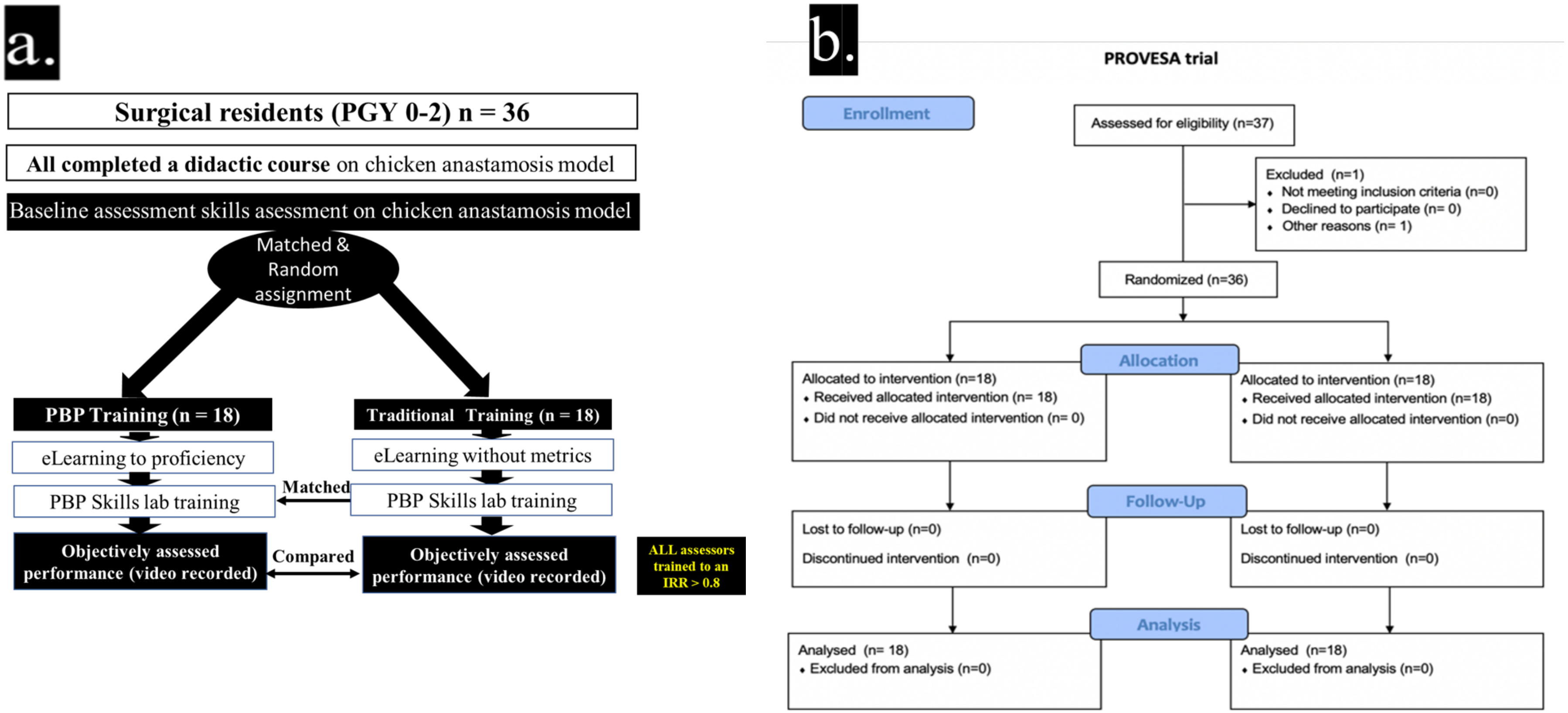

Figure 1a and b. a) The design of the multicentre, prospective, randomised, matched and blinded study and b) The CONSORT 2010 flow diagram of the PROVESA trial.

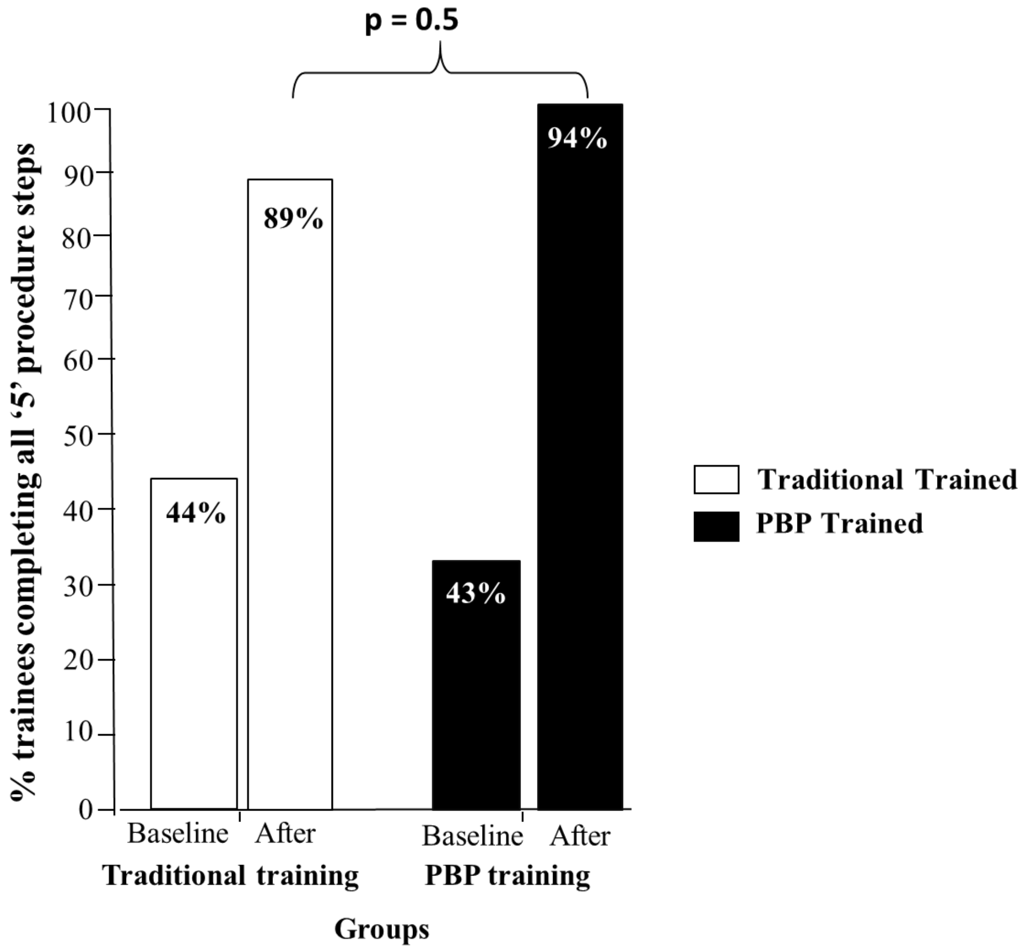

Figure 2. The percentage of trainees who completed all five procedure steps at baseline assessment and at the end of their training in the video-recorded and blinded assessment.

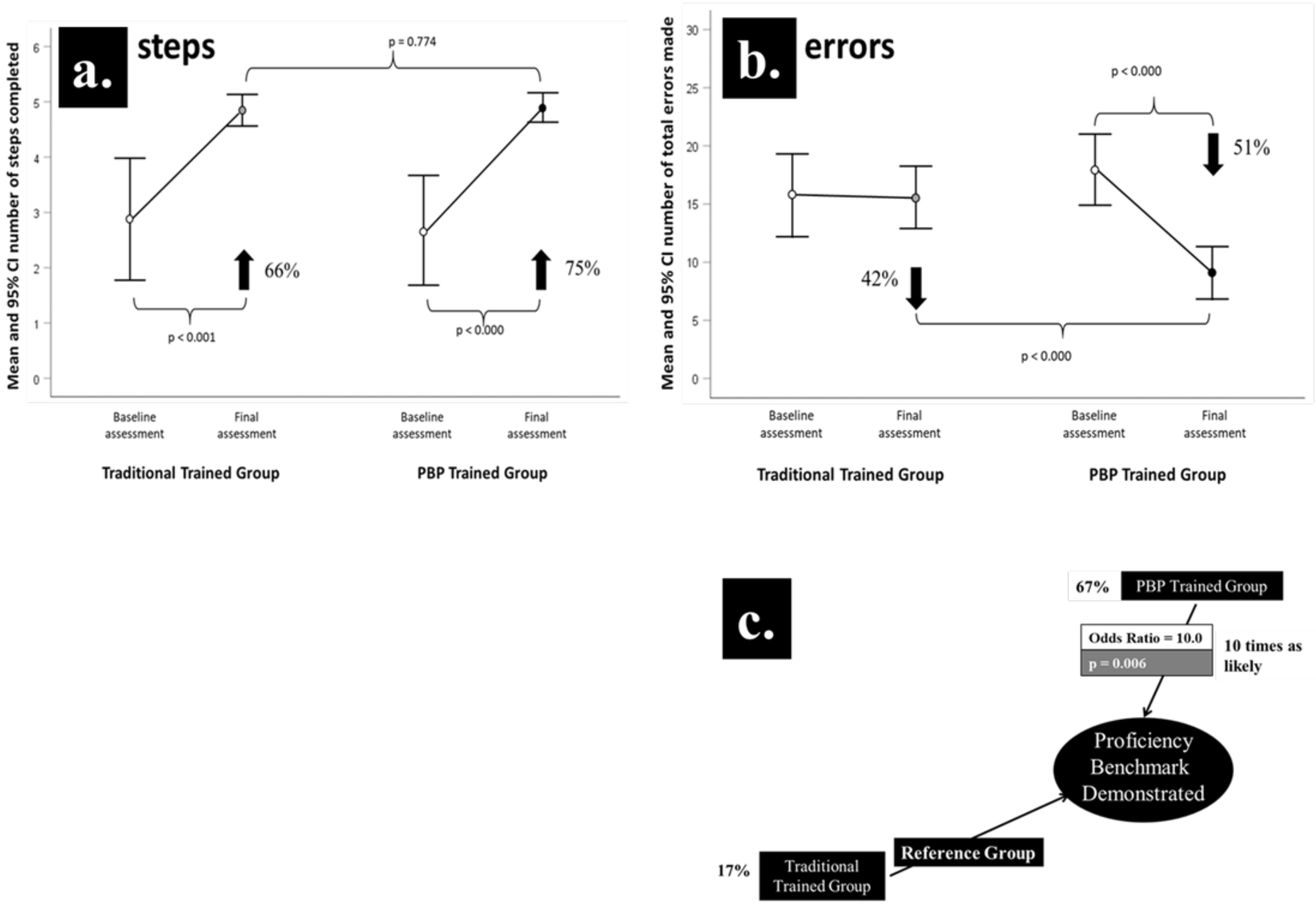

Figure 3a - c. The mean and 95% confidence intervals (CI) of a.) procedure steps completed and b.) the number of errors made by the Traditional and PBP trained Groups at Baseline and at the end of training and c.) Logistic Regression Analysis for the relative differences between the Traditional Trained Group (Reference Group for comparison) and the PBP trained group proficiency demonstration at the end of training.