Use of Robotic Anastomosis Competency Evaluation (RACE) tool for assessment of surgical competency during urethrovesical anastomosis

Hijab Khan[1]; Justen D. Kozlowski[1]; Ahmed A. Hussein, MD[1,2]; Mohamed Sharif[1]; Youssef Ahmed, MD[1]; Paul May[1]; Yana Hammond[1]; Kevin Stone[1]; Basim Ahmad[1]; Adam Cole[1]; Adam Hasasneh[1]; Sana Raheem[1]; Khurshid A. Guru, MD[1]

[1]Roswell Park Cancer Institute, Buffalo, NY, United States; [2]Cairo University, Egypt

Cite as: *Can Urol Assoc J* 2019;13(1):E10-6. http://dx.doi.org/10.5489/cuaj.5348

Published online July 24, 2018

Abstract

Introduction: We sought to evaluate the Robotic Anastomosis Competency Evaluation (RACE) — a validated tool that objectively quantifies surgical skills specifically for urethrovesical anastomosis (UVA) — as a means to track progress of trainees, as well as to determine its predictive value.

Methods: UVAs performed by trainees at our institution were evaluated using RACE over a period of two years. Trainees were supervised by an experienced robotic surgeon. Outcomes included trainee-related variables (RACE score, proportion of UVA performed by trainee, and suturing speed), and clinical outcomes (total UVA duration, postoperative urinary continence, and UVA-related complications). Significance was determined using linear regression analysis.

Results: A total of 51 UVAs performed by six trainees were evaluated. Trainee RACE scores (19.8 to 22.3; p=0.01) and trainee proportion of UVA (67% to 80%; p=0.003) improved significantly over time. Trainee suture speed was significantly associated with RACE score (mean speed range 0.54–0.74 sutures/minute; p=0.03). Neither urinary continence at six weeks nor at six months was significantly associated with RACE score (p=0.17 and p=0.15, respectively), and only one UVA-related postoperative complication was reported.

Conclusions: Trainee RACE scores improved and the proportion of UVA performed by trainees increased over time. RACE can be used as an objective measure of surgical performance during training. Strict mentor supervision allowed safe training without compromising patient outcomes.

Introduction

Surgical skills and technical proficiency impact the quality of life and oncological outcomes of patients independent of case volume and surgeon experience.[1] Robot-assisted surgery offers several advantages for patients in terms of reduced blood loss, fewer blood transfusions, improved convalescence, and quicker recovery; and for surgeons in terms of improved dexterity, facilitated suturing, access to deep anatomical areas (such as the pelvis), magnification, and 3-D visualization. However, teaching surgical skills may become more challenging, as a new set of surgical and non-technical skills, as well as familiarity with the human-robot interface, are required. Residency and fellowship training programs face the challenge of transferring surgical expertise from mentors to trainees without jeopardizing patient outcomes. Additionally, a standardized definition of what constitutes surgical proficiency is subject to variation, which results in a lack of consistent feedback to trainees. Metrics, such as number of procedures performed or console hours, have been previously suggested, but none of these provide objective or real-time feedback of surgical proficiency. Another challenge is that most surgical teaching occurs in the operating room (OR), which can significantly increase operative time and cost, and may potentially impact patient outcomes.[2] The American Urological Association (AUA) recommended development of validated tools that can be used for objective evaluation of robotic surgical training, and provide structured feedback for trainees. Such tools can also be valuable in surgeon credentialing and remediation.[3]

Despite the paramount increase in the use of robot-assisted surgery, there is still lack of a standardized definition of robotic surgical proficiency. Several tools, like the Global Operative Assessment of Laparoscopic Skills (GOALS) and the Global Evaluative Assessment of Robotic Skills (GEARS), have been developed to objectively assess laparoscopic and robotic skills.[4,5] Both tools have been shown very useful in terms of assessing generic laparoscopic and robotic skills, but they lack procedure-specific measures.

Urethrovesical anastomosis (UVA) is a critical step during robot-assisted radical prostatectomy (RARP). Technical mishaps can result in considerable morbidity, including anastomotic leaks, urinary incontinence, and stricture formation. UVA represents an ideal surgical step to assess and track suturing proficiency. Our group has previously developed and validated the Robotic Anastomosis Competency

Evaluation (RACE) to objectively quantify surgical performance during UVA and provide structured feedback to trainees.[6] RACE analyzes UVAs in six domains using well-defined anchors with specific descriptions that include supplementary illustrations for clarity and minimize variation among raters (Supplementary Fig. 1). In this study, we sought to evaluate the predictive validity of RACE. We hypothesized that RACE scores would improve with time. We also investigated association of RACE with clinical outcomes.

Methods

As part of a training quality initiative at Roswell Park Comprehensive Cancer Center, trainee UVA performances during RARPs were evaluated over a period of two years (Institutional Review Board: BDR 062715). All trainees present during this time were eligible for the study. Trainees included chief residents and fellows strictly supervised and assessed by one fellowship-trained urologist (experience with >1200 RARPs and 6000 console hours) who was involved in the development of the RACE tool and was thus familiar with its use. No trainees had prior laparoscopic prostatectomy experience, although some had prior RARP experience. The number of evaluations and the length of evaluation periods varied by trainee (Supplementary Table 1).

Outcomes included both trainee-related and clinical outcomes. Trainee-related outcomes included RACE scores, which were obtained during real-time assessment of trainees performing UVA, as well as UVA proportion performed by trainee and trainee suturing speed, which were determined in a video review of the surgery. Clinical outcomes included total UVA duration, UVA-related postoperative complications, and urinary continence at six weeks and six months.

UVA technique

Two V-loc sutures tied together were used for UVA. For the posterior plate, the first suture was started in the bladder neck from outside-in at the six o'clock position, then into the urethra from inside-out at six o'clock. A running suture was then completed on both sides until both met in the 12 o'clock position, where the two sutures were tied together. Fellows usually start with the anterior plate and, as they progress through the fellowship, start performing more of the posterior plate.

RACE scores

Six domains were assessed: needle positioning, needle entry, needle driving and tissue trauma, suture placement, tissue approximation, and knot tying. Each domain was scored on a 1–5 Likert scale tabulated to a final score out of 30, where specific descriptions were provided for scores 1 (the worst performance), 3 (intermediate), and 5 (the ideal performance). Mentoring surgeons completed hard copy RACE forms during trainees' real-time performances of the UVA.[6]

Trainee-related outcomes

Trainee-related outcomes were determined by retrospectively reviewing video and audio recordings of observed procedures to delineate who was on the console. Our methodology for OR environment recording has been previously published.[7] Total UVA duration was measured from the first needle insertion into the bladder to when the final knot was tied. Trainee UVA duration was measured when a switch from mentor to trainee took place, starting at the trainee's first robotic movement and ending when control was relinquished. UVA proportion performed by trainee was calculated using the proportion of trainee UVA duration to total UVA duration. Suture speed was calculated by counting suture loops made by trainees divided by the trainee UVA duration.

Clinical outcomes

Our prospectively maintained quality assurance database was queried for patient demographics; disease characteristics; postoperative outcomes, including complications specific to UVA (urinary leakage, urinary ascites, bladder neck contracture, and urethral stricture); and urinary continence at six weeks and six months. Patient continence was defined as the number of self-reported absorbent pads used by a patient per day.

Statistical analysis

Descriptive statistics were used to illustrate the data. Linear regression analysis was used to determine effect size and significance for associations between the trainee evaluation variables, such as UVA proportion, suture speed, and RACE score, with time from first assessment. Linear regression analysis was also used to determine the associations between RACE score and other trainee evaluation variables to determine the score's predictive value. Statistical significance was set at alpha level 0.05. All statistical analyses were performed using R Core Team and SAS, version 9.4.

Results

The final cohort comprised of six trainees (two fellows were evaluated only once and were excluded from the final analysis, as the majority of our outcomes depended on performance over time). All UVAs performed during the study period and involving trainee participation were included in the analysis (n=51). Mean age of patients was 60 years, 26 (50%) had prior abdominal surgery, and 13 (25%) had American Society of Anesthesiologists (ASA) score ≥3 (Supplementary Table 2). Trainees began their fellowship by performing only

the anterior portion of the UVA before eventually advancing to performing portions of the posterior plate of the UVA. As experience was acquired, trainees performed the entire UVA without mentor interference (n=10). Trainee RACE scores improved significantly over time (p=0.01) (Fig. 1A). The trainee proportion of the UVA increased significantly with time (from 67% to 80%; p=0.003) (Fig. 1B). Trainee suture speed was significantly associated with RACE scores, with mean speed increasing from 0.54 to 0.74 sutures/minute (p=0.03) (Fig. 2). Total UVA duration was neither significantly associated with time (p=0.86), nor with RACE score (p=0.17) (Fig. 3). Neither urinary continence at six weeks nor six months was significantly associated with RACE score (p=0.17 and p=0.15, respectively) (Fig. 4). An anastomotic leak was the single UVA-related postoperative complication reported.

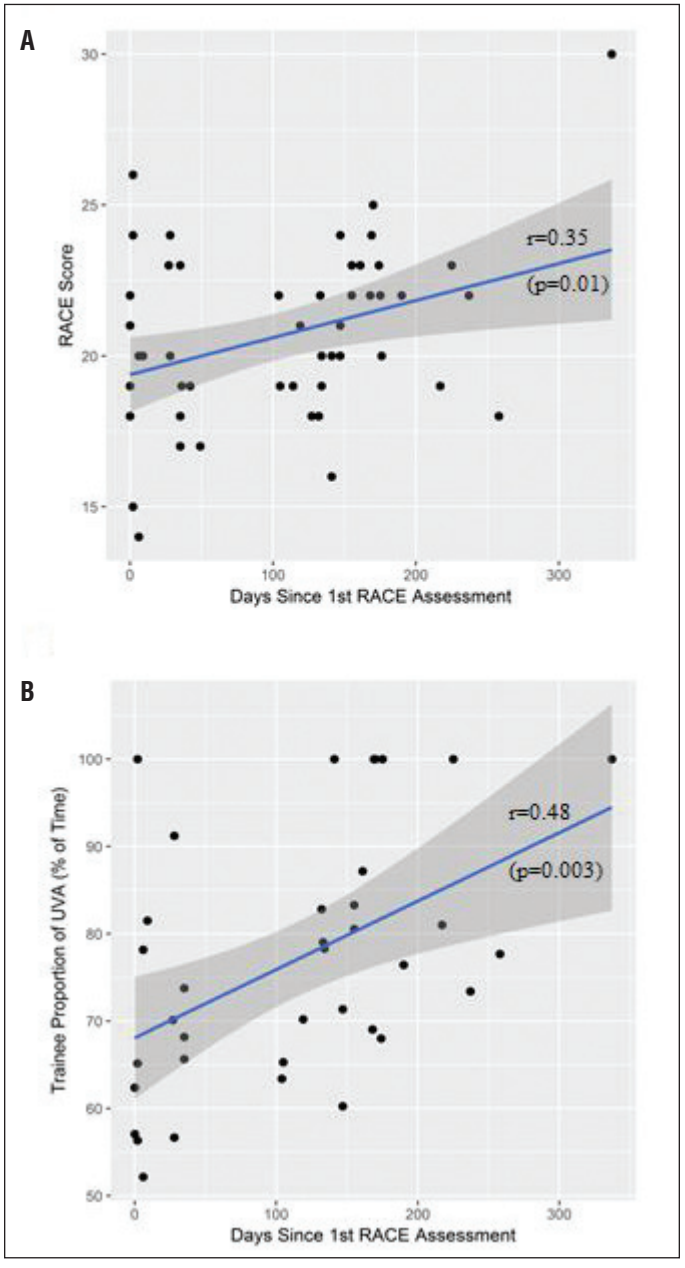


Fig. 1. (A) Change in Robotic Anastomosis Competency Evaluation (RACE) scores over time. ***(B)*** Trainee urethrovesical anastomosis (UVA) proportion over time.

Discussion

Understanding trainee progress and the transfer of skills between mentor and trainee is critical in the field of robot-assisted surgery. Modular training offers discrete and structured steps for trainees to complete as they learn the skills needed for robot-assisted surgery without compromising clinical outcomes. This begins with preclinical training, including understanding of laparoscopy, basic skills practice, familiarity with the da Vinci Surgical System®, and surgical practice on animate models. The novel presence of a console in robot-assisted surgery changes the dynamic of mentor-trainee interactions. One study examining EEG data found that the mental workload of a mentor observing a trainee's surgical performance increased when the mentor perceived the trainee as performing poorly.[8] Another found that mentor's brain functional connectivity was higher when observing low-quality trainee performances compared to high-quality ones.[9] Stress during surgery could compromise patient outcomes, and Shafiei et al demonstrated that mentors can face stress due to observing low-quality trainee performance and worrying about proficiency.[10]

Preclinical training is followed by clinical training, during which trainees observe and then gradually participate, in a

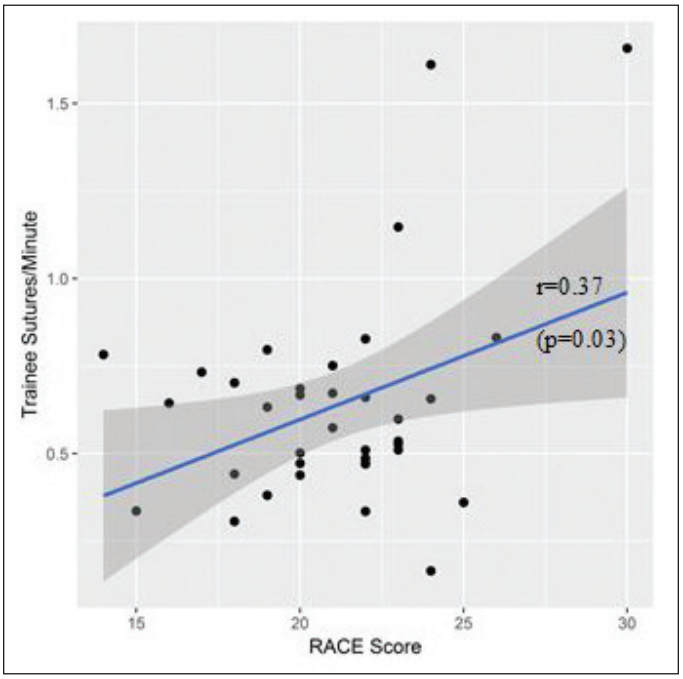


Fig. 2. Robotic Anastomosis Competency Evaluation (RACE) score and trainee suture speed.

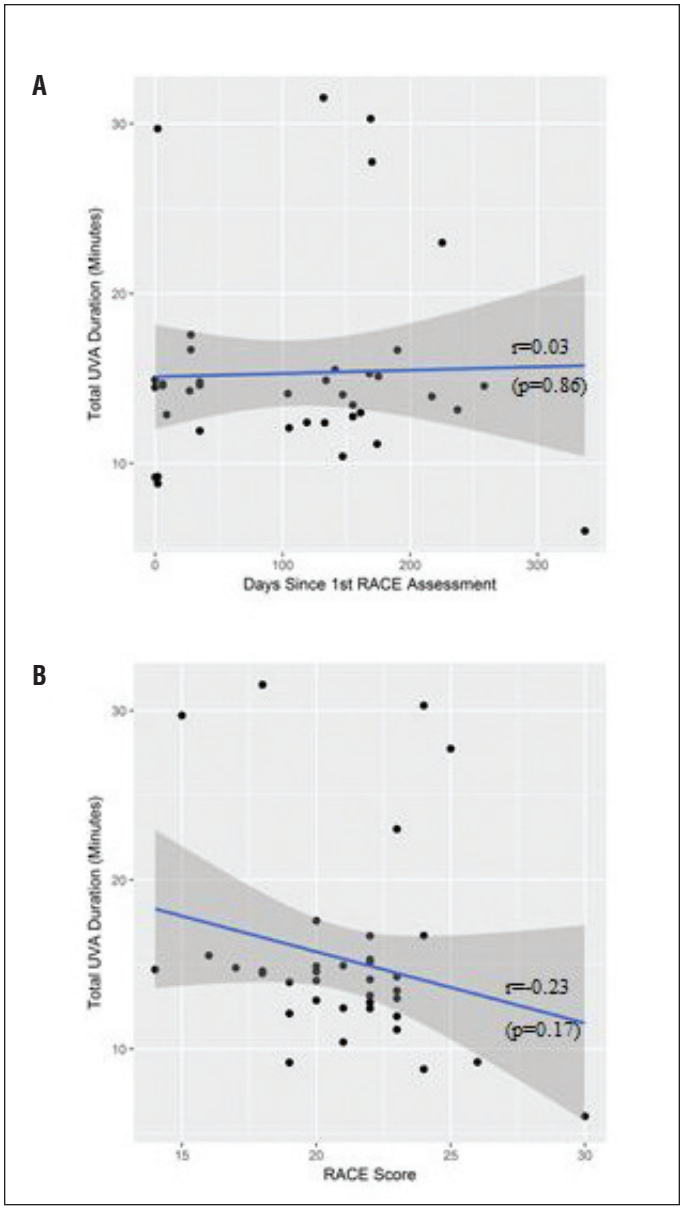


Fig. 3. (A) Total urethrovesical anastomosis (UVA) duration over time. ***(B)*** Total UVA duration and Robotic Anastomosis Competency Evaluation (RACE) score.

stepwise fashion, in surgical procedures.[3] Cognitive training involves improving and tracking abilities like memory, attention, situational awareness, and skill acquisition relevant to robot-assisted surgery. Methods include virtual reality and haptics, technical training, and mentorship. Brain functional states have been used to objectively evaluate skill acquisition.[11] Dreyfus' model of skill acquisition describes a five-stage learning process, from requiring monitoring and feedback to performing skills without conscious attention.[12] Fitts and Posner's model of motor skill acquisition suggests that trainees move through three stages of learning: cognitive, associative, and autonomous, becoming more competent and requiring less conscious attention as they progress.[13]

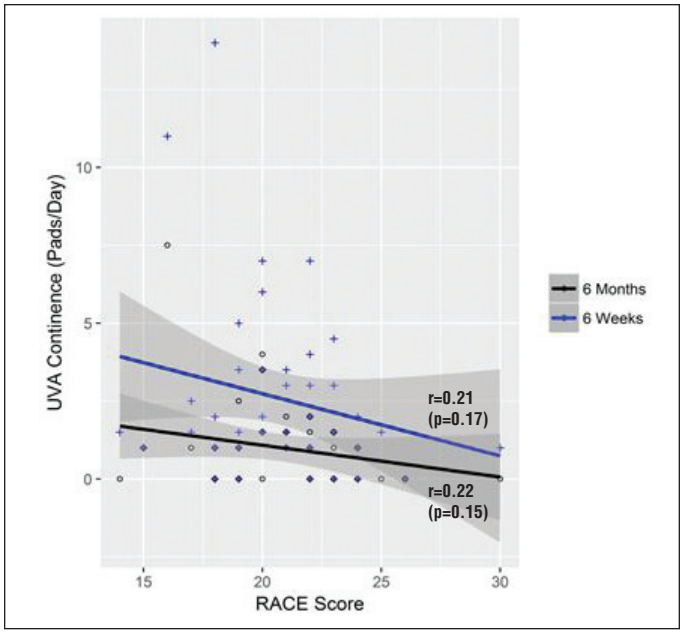


Fig. 4. Robotic Anastomosis Competency Evaluation (RACE) score and continence at six weeks and six months

Despite the paramount increase in the use of robot-assisted surgery, there is still lack of a standardized definition of surgical proficiency. Several tools, like GOALS and GEARS have been developed to objectively assess laparoscopic and robotic skills.[4,5] Both tools have shown very useful in terms of assessing generic laparoscopic and robotic skills, but they lack procedure-specific measures.

To monitor progress in surgical training, assessments for surgical procedures have been developed and validated, such as the RACE assessment used in this study. Prior assessment tools using similar methodology have been developed and validated, including the Pelvic Lymphadenectomy Assessment and Completion Evaluation (PLACE), Prostatectomy Assessment and Competency Evaluation (PACE), Robotic Hysterectomy Assessment Score (RHAS), and Cystectomy Assessment and Surgical Evaluation (CASE).[14-17] PACE and CASE deconstructed their respective procedures into multiple domains assessed on a five-point Likert scale, built and validated by experts. PLACE divided pelvic lymph node dissection (PLND) into three zones as determined by experts. Each zone was accompanied by an illustration and raters used the zones to score completeness of PLND. Lovegrove et al developed and validated a tool that both evaluated technical skills and constructed learning curves for the procedural steps of RARP.[18] This tool requires the mentor to assess 41 subprocesses, but remains subject to variation because it does not offer descriptive anchors for performance scores.

Our study confirmed that RACE is a good predictor of trainee performance and mastery of skill using the da Vinci Surgical System®. The significant increase in trainee RACE scores over time likely indicates the trainee's improvement

in the RACE domains, while the significant linear relationship between time and trainee proportion of the UVA suggests gradual trainee inclusion in the surgical procedure. There is controversy about the possible effects of trainees' participation on surgical outcomes. Some literature has suggested that trainee inclusion in procedures can affect clinical outcomes through increased time in the OR and more complications, while others have found no such effect.[2,19,20] Our study found no significant association between UVA duration and time since first RACE assessment, or between UVA duration and RACE score.

The single UVA-related complication, and the lack of significant association between RACE score and continence, suggests trainee inclusion and performance did not compromise operative outcomes. However, other causes, like the smaller number of procedures included in the study or unmeasured surgical components, including bladder neck dissection, neurovascular bundle, and pelvic floor muscles preservation, could affect outcomes. Outcomes may also reflect timely mentor intervention in surgery, mitigating adverse effect on perioperative outcomes. The trainees' prior simulation training, which shortens the learning curve for skill acquisition, could have strengthened their performance.[3] Perhaps their gradual inclusion over time allowed for mastery of skill without the pressure of performing an entire procedure. Mastery of skill requires the ability to complete a procedure as though it was second nature, while simultaneously being prepared to consciously focus on the procedure in the event of an unusual situation.[21] Trainees move through the cognitive, associative, and autonomous phases before they master a skill. The cognitive phase, which consists of intellectualized, consciously focused task performance, is best practiced outside of the OR, which allows for more complex learning during surgery.[22]

RACE and similar assessments are valuable in that they break down procedures into distinct steps. This facilitates structured and tailored feedback specific to the areas of surgery in which the trainee needs practice. For example, if a trainee scores poorly on knot-tying in a UVA, he/she can undergo remedial training to improve that particular skill. The RACE assessment also allows for more objective feedback of trainees who are learning from multiple mentors, as the validated tool was shown to demonstrate inter-rater reliability.[6] Such objective feedback can be incorporated into institutional processes for credentialing.

This study has a number of limitations. RACE assessment looks only at UVA, a single domain of RARP. The PACE assessment may be used to evaluate the full RARP procedure.[16] Only one rater was used to assess the trainees. As this data was gathered during a quality assurance procedure at our institution, the study was not blinded and trainees were evaluated in real time. This may have allowed for bias in the single reviewer's assessment of the trainees. The varying length of time spent by trainees in fellowship and residency prior to evaluation may have influenced their UVA performance and could account for the range in baseline RACE scores. However, this has minimal implications for our analysis, which assesses the changes in RACE scores over time and primarily considers trends as opposed to individual performances. We considered the trainee proportion of the UVA solely as a trainee-related outcome. It could also be affected by the mentors' confidence in the trainee's ability, and their assessment of the surgical case's complexity. However, as the mentor becomes more comfortable with the trainee's performance over time, they may allow the trainee to perform more of the UVA and include them in more complex cases. Additionally, the association of continence and RACE score may be affected by factors like patient age, prior continence status, and the quality of the apical dissection. The continence data itself was gathered through self-reported information from patients, which could be subject to recall bias. Nevertheless, this is part of our institution's routine assessment of patients who undergo RARP, and this information is clearly documented in the patients' charts. Finally, the small number of procedures assessed and the single institutional study design inherently limits the generalizability of results.

Conclusion

RACE predicted improvement of trainee surgical skill in UVA over time. It can be used as an objective measure of UVA performance. Patient outcomes may not be compromised by the strictly supervised inclusion of trainees in surgical procedures.

Competing interests: The authors report no competing personal or financial interest related to this work.

Acknowledgment: This study was supported by the Roswell Park Alliance Foundation.

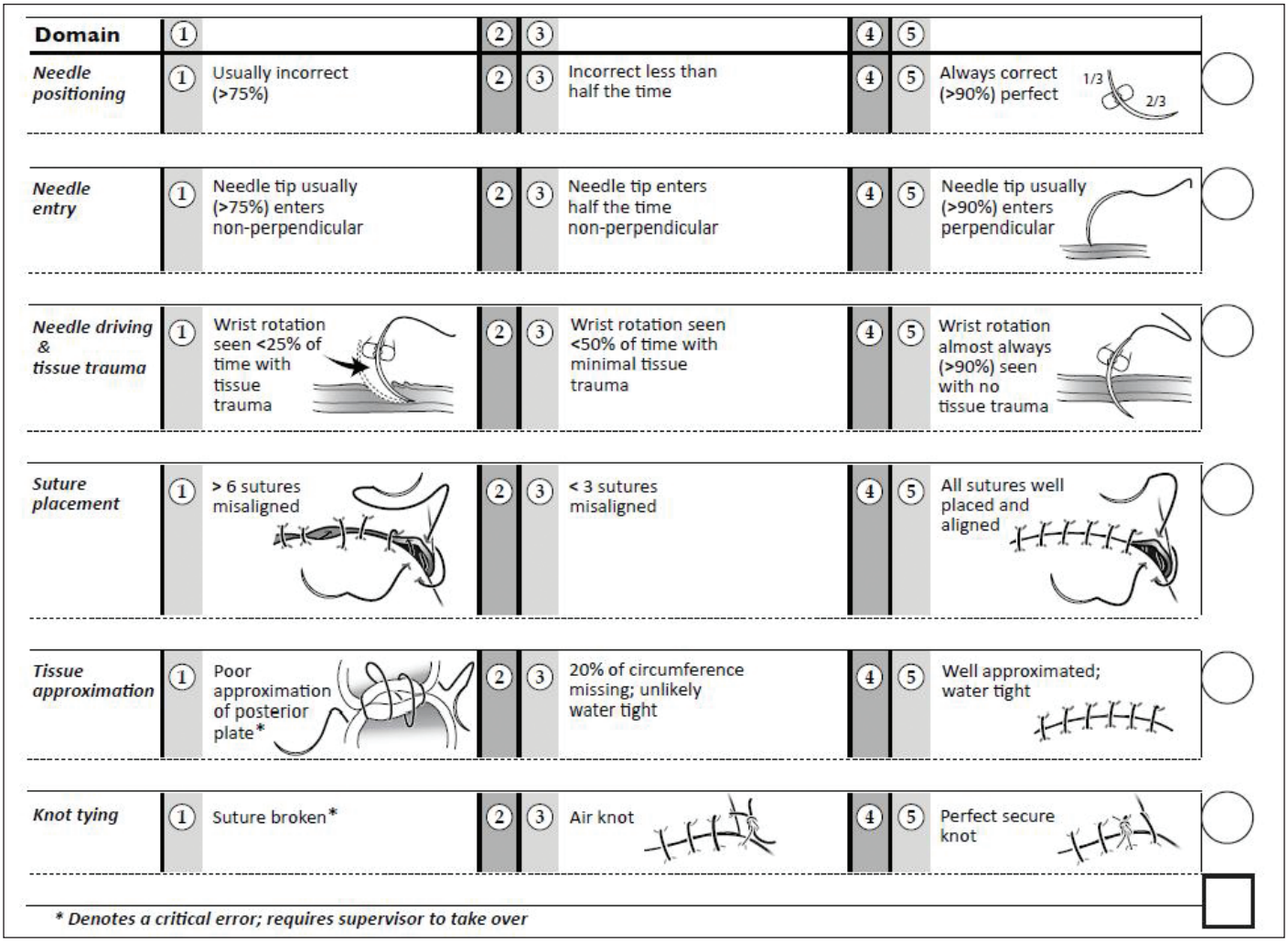
This paper has been peer-reviewed.

References

1. Vickers A, Savage C, Bianco F, et al. Cancer control and functional outcomes after radical prostatectomy as markers of surgical quality: Analysis of heterogeneity between surgeons at a single cancer centre. *Eur Urol* 2011;59:317-22. https://doi.org/10.1016/j.eururo.2010.10.045
2. Babineau TJ, Becker J, Gibbons G, et al. The «cost» of operative training for surgical residents. *Arch Surg* 2004;139:366-9. https://doi.org/10.1001/archsurg.139.4.366
3. Lee JY, Mucksavage P, Sundaram CP, et al. Best practices for robotic surgery training and credentialing. *J Urol* 2011;185:1191-7. https://doi.org/10.1016/j.juro.2010.11.067
4. Vassiliou MC, Feldman LS, Andrew CG, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg* 2005;190:107-13. https://doi.org/10.1016/j.amjsurg.2005.04.004
5. Goh AC, Goldfarb DW, Sander JC, et al. Global evaluative assessment of robotic skills: Validation of a clinical assessment tool to measure robotic surgical skills. *J Urol* 2012;187:247-52. https://doi.org/10.1016/j.juro.2011.09.032

6. Raza SJ, Field E, Jay C, et al. Surgical competency for urethrovesical anastomosis during robot-assisted radical prostatectomy: Development and validation of the robotic anastomosis competency evaluation. *Urology* 2015;85:27-32. https://doi.org/10.1016/j.urology.2014.09.017
7. Tiferes J, Hussein AA, Bisantz A, et al. The loud surgeon behind the console: Understanding team activities during robot-assisted surgery. *J Surg Educ* 2016;73:504-12. https://doi.org/10.1016/j.jsurg.2015.12.009
8. Hussein AA, Shafiei SB, Sharif M, et al. Technical mentorship during robot-assisted surgery: A cognitive analysis. *BJU Int* 2016;118:429-36. https://doi.org/10.1111/bju.13445
9. Shafiei SB, Hussein AA, Kozlowski JD, et al. PD46-02 Looking for your own reflection: Assessing brain functional state of surgical mentor during robot-assisted surgery. *J Urol* 2017;197:e890. https://doi.org/10.1016/j.juro.2017.02.2374
10. Shafiei SB, Hussein AA, Ahmed Y, et al. MP51-05 Does trainee performance impact surgeon's stress during robot-assisted surgery? *J Urol* 2017;197:e695. https://doi.org/10.1016/j.juro.2017.02.1614
11. Shafiei SB, Hussein AA, Guru KA. Cognitive learning and its future in urology: Surgical skills teaching and assessment. *Curr Opin Urol* 2017;27:342-7. https://doi.org/10.1097/MOU.0000000000000408
12. Dreyfus S, Dreyfus H. A five-stage model of the mental activities involved in directed skills acquisition. In. University of California, Berkeley1980. https://doi.org/10.21236/ADA084551
13. Sadideen H, Kneebone R. Practical skills teaching in contemporary surgical education: How can educational theory be applied to promote effective learning? *Am J Surg* 2012;204:396-401. https://doi.org/10.1016/j.amjsurg.2011.12.020
14. Frederick PJ, Szender JB, Hussein AA, et al. Surgical competency for robot-assisted hysterectomy: Development and validation of a robotic hysterectomy assessment score (RHAS). *J Minim Invasive Gynecol* 2017;24:55-61. https://doi.org/10.1016/j.jmig.2016.10.004
15. Hussein AA, Hinata N, Dibaj S, et al. Development, validation and clinical application of pelvic lymphadenectomy assessment and completion evaluation: Intraoperative assessment of lymph node dissection after robot-assisted radical cystectomy for bladder cancer. *BJU Int* 2017;119:879-84. https://doi.org/10.1111/bju.13748
16. Hussein AA, Ahmed YE, Kozlowski JD, et al. Robot-assisted approach to 'W'-configuration urinary diversion: A step-by-step technique. *BJU Int* 2017;120:152-7. https://doi.org/10.1111/bju.13824
17. Hussein AA, Sexton K, Maxwell M, et al. MP34-10 development and validation of cystectomy assessment and surgical evaluation (CASE) scoring for male radical cystectomy. *J Urol* 2017;197.
18. Lovegrove C, Novara G, Mottrie A, et al. Structured and modular training pathway for robot-assisted radical prostatectomy (RARP): Validation of the RARP assessment score and learning curve assessment. *Eur Urol* 2016;69:526-35. https://doi.org/10.1016/j.eururo.2015.10.048
19. Goodwin AT, Birdi I, Ramesh TP, et al. Effect of surgical training on outcome and hospital costs in coronary surgery. *Heart* 2001;85:454-7. https://doi.org/10.1136/heart.85.4.454
20. Schroeck FR, Palha de Sousa CA, Kalman RA, et al. **Trainees do not negatively impact the institutional learning** curve for robotic prostatectomy as characterized by operative time, estimated blood loss, and positive surgical margin rate. *Urology* 2008;71:597-601. https://doi.org/10.1016/j.urology.2007.12.023
21. Guru KA, Shafiei SB, Khan A, et al. Understanding cognitive performance during robot-assisted surgery. *Urology* 2015;86:751-7. https://doi.org/10.1016/j.urology.2015.07.028
22. Reznick RK, MacRae H. Teaching surgical skills — changes in the wind. *N Engl J Med* 2006;355:2664-9. https://doi.org/10.1056/NEJMra054785

Correspondence: Dr. Ahmed A. Hussein, Cairo University, Egypt; ahmedalyhussein@kasralainy.edu.eg



Supplementary Fig. 1. Robotic Anastomosis Competency Evaluation (RACE) form.

Supplementary Table 1. Trainee characteristics

Trainee	Prior robotic experience	Assessments (n)	Length of evaluation period (days)
1	Yes	7	36
2	No	5	237
3	Yes	8	147
4	No	5	258
5	No	15	170
6	No	11	337

Supplementary Table 2. Characteristics of patients who underwent RARP

Variable	n (%)
Number of patients	51
Age, years mean (SD)	60 (6)
BMI, kg/m^2, mean (SD)	30.2 (4)
ASA score, median, (IQR)	2 (2–2.5)
ASA score ≥3	13 (25)
Prior surgery	26 (51)
Operative time, minutes, median (IQR)	174 (151–189)
EBL >500 ml	2 (4)
≥pT3	20 (39)
pN positive	1 (2)
PSM	13 (25)
Lymph node count, mean (SD)	6 (5.9)
Complications	16 (31)
30-day complications	10 (20)
30–90-day complications	6 (12)
High grade complications	3 (6)
UVA-related complications	1 (2)

ASA: American Society of Anesthesiologists; BMI: body mass index; EBL: estimated blood loss; IQR: interquartile range; pN: node positive; PSM: positive surgical margins; RARP: robot-assisted radical prostatectomy; SD: standard deviation; pT3: pathologic stage T3; UVA: urethrovesical anastomosis.